AFIT/GOR/ENS/93M-07

MODIFIED ANDERSON-DARLING AND CRAMER-VON MISES GOODNESS-OF-FIT TESTS FOR THE NORMAL DISTRIBUTION

THESIS

David Alan Gwinn, Sr.
Captain, USAF

AFIT/GOR/ENS/93M-07

AFIT/GOR/ENS/93M-07

MODIFIED ANDERSON-DARLING AND CRAMER-VON MISES GOODNESS-OF-FIT TESTS FOR THE NORMAL DISTRIBUTION

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Operations Research

David Alan Gwinn, Sr., B.A.
Captain, USAF

March, 1993

Approved for public release; distribution unlimited

THESIS APPROVAL

STUDENT: Capt David A. Gwinn, Sr. CLASS: GOR-93M

THESIS TITLE: MODIFIED ANDERSON-DARLING AND CRAMER-VON MISES GOODNESS-OF-FIT TESTS FOR THE NORMAL DISTRIBUTION

DEFENSE DATE: March 8, 1993

COMMITTEE:	NAME/DEPARTMENT	SIGNATURE
Advisor	Dr. Albert H. Moore/ENC	Albert H Moore
Reader	Dr. Joseph P. Cain/ENS	Joseph P Cain

Preface

This thesis develops new, modified Anderson-Darling and Cramer-von Mises goodness-of-fit tests for normal distributions with parameters estimated from the sample. The complete critical value tables are presented for each test. These tables can be used to test whether a set of observed values follows a normal distribution. Additionally, the power tables are presented for each distribution to help the readers understand the power relations clearly with different sample sizes.

I wish to thank Dr. Albert H. Moore, my advisor for this research topic, his guidance and his never-ending support. I would also like to thank Dr. Joseph Cain, my reader for his assistance. Finally, I wish to thank my lovely wife, Jerri, for her understanding and patience during the course of my work away from her.

David Alan Gwinn, Sr.

List of Figures

List of Tables

Table		Page

Abstract

New techniques for calculating goodness-of-fit statistics for normal distributions with parameters estimated from the sample are investigated. Samples are generated for a Normal(0,1) distribution. Critical values are calculated for five modifications to the Anderson-Darling statistic and five modifications to the Cramer-Von Mises statistic. An extensive power study is done to test the power of the new statistics versus the power of the unmodified statistics.

Powers of six of the new statistics show minimal to no improvement, two of the new statistics show a marked decrease in power, and two of the new statistics show an overall increase in power over the unmodified statistics. One of these two improved statistics was the obvious better of the two, and it was a modification to the Anderson-Darling statistic.

Complete tables of critical values for sample sizes n=4 through n=50 are included for all Anderson-Darling and Cramer-Von Mises statistics, both modified and unmodified.

MODIFIED ANDERSON-DARLING AND CRAMER-VON MISES GOODNESS-OF-FIT TESTS FOR THE NORMAL DISTRIBUTION

I. Introduction

1.1 Background

The amount of money the Air Force receives from Congress to accomplish the mission continues to shrink. With technological advances, problem solving often becomes more and more complex. It is not enough to just solve the problems, it is essential that they be solved as efficiently and effectively as possible. When these problems involve quantifiable data on which analysis can be accomplished, decision-makers rely on analysts to assist them. One problem solving technique often used by analysts to help decision-makers is simulation. Simulation is a modelling of the real-world that incorporates the critical characteristics of the real-world problem. Simulation is much less expensive than conducting live tests and can be modified as the problem changes. Decision-makers use the results of the analyst's simulation to make efficient and effective decisions.

Simulation models are made up of many different features and parts. One of these features is that they mimic at least one real-world random phenomena. For example, the random arrival of customers at an Air Force Accounting and Finance office would be a feature in a simulation used to examine customer service times. In order for the simulation to properly mimic the real-world random phenomena, a sample of real-world data must be collected and its characteristics scrutinized. (21:43). After collecting the data, the analysts compare this data to a theorized cumulative distribution function, i.e. a sample is estimated to be from the exponential

distribution with $\lambda = 10$. At this point, the analysts want to know how well the data fits the hypothesized distribution. To check if the data fits the hypothesized distribution, a test called a goodness-of-fit test (GOFT) is performed. GOFTs are a class of statistical comparisons between the characteristics of a sample of data and the characteristics of a estimated theoretical distribution with the hope that the two match.

A GOFT is used to find the appropriate distribution of the sample data so that assumptions and predictions can be made about the behavior of the population. This, in turn allows claims to be made about the population and subsequently, about the validity of the simulation. For example, "With 95 percent confidence, I predict this missile will destroy the target."

A GOFT is conducted by going through these four steps.

1. Collect a sample of data for the event of interest.
2. Estimate a distribution of this sample data.
3. Choose the type of GOFT to use and compute the appropriate test statistic.
4. Compare this test statistic to the critical value from statistical tables.

If the test statistic is less than or equal to the table value, then the estimated distribution is not rejected. If the test statistic is greater than the table value, then the estimated distribution is rejected and the GOFT is repeated with another estimate of the distribution. When a test shows a good fit, it gives a degree of confidence to the analyst that the hypothesized distribution describes the real-world data and that it can be used effectively in a simulation model or any other analysis. Accuracy is critical at this stage of the game; no matter how sophisticated the simulation model is, the choice of an incorrect distribution for a random process can make the ensuing analysis useless (17).

There is not just one test that will tell the analyst whether or not an estimated distribution is correct. Over the years, many different types of GOFTs have been

developed. With the development of each new GOFT, the goal was to create a GOFT that was better than the last. The most commonly used GOFTs are the Chi-square, the Kolmogorov-Smirnov (KS), the Anderson-Darling (A-D) and the Cramer-von Mises (C-VM) (7:109-114). Many of the more recent developments with GOFTs have been via minor modifications to these GOFTs.

1.2 Problem Statement

The Anderson-Darling (A-D) and Cramer-von Mises (C-VM) GOFT statistics have been modified by earlier researchers to be more powerful for the normal distribution; however, I believe that there is room for improvement. My research goal is to modify the A-D and C-VM tests to produce a better GOFT for normality with the parameters estimated from the sample.

1.3 Methodology

The research effort will consist of the following steps:

1. Produce unmodified A-D and C-VM critical value tables for the normal distribution.
2. Produce modified A-D and C-VM critical value tables for the normal distribution.
3. Conduct tests to compare the powers of these modified and unmodified statistics.

1.3.1 Critical Value Tables. In the first step, critical value tables will be generated for the unmodified A-D and C-VM statistics. These tables will be developed by generating random samples from a normal distribution via Monte Carlo simulation using samples of size four.

Then using the appropriate equations outlined in chapter three, the A-D and C-VM statistics will be calculated. These steps will be repeated 10,000 times, using

a sample size of four, to generate 10,000 independent A-D statistics and 10,000 independent C-VM statistics. These statistics will then be used to interpolate the critical values for the respective tables. Once this is accomplished, it will then be repeated for samples of size n = 5,6,...,50.

This same sequence of events will be accomplished using the modified A-D and C-VM statistical values. Which will in turn allow step three to be worked as we will have critical value tables for both the unmodified and the modified A-D and C-VM tests.

1.3.2 Power Comparison. The power of a statistical test is the probability of correctly rejecting a false null hypothesis. The gist of the power study is in testing to see if the null hypothesis is rejected when in fact it should be rejected; i.e., what percentage of the time is the null, that I have a normal sample, rejected when in fact my sample is from an exponential. The higher the percentage, the better, more powerful the test. These powers of the modified and unmodified A-D and C-VM tests will be compared to determine which amongst them can best detect whether or not a sample of data comes from a population characterized by the normal distribution.

To perform these power tests, the A-D and C-VM test statistics will be calculated under the null hypothesis that the random samples follow the normal distribution with the specified parameters. The calculated statistics for each test will then be compared to the appropriate row and column of the corresponding critical value tables obtained in steps one and step two to determine whether to reject the null hypothesis. This will be repeated 10,000 times. Then, the number of times each statistic exceeds the respective critical value will be divided by 10,000 and this will give us the power of the test - the percentage of the time the null hypothesis is correctly rejected. This power computation will be performed with data created from Monte Carlo simulation for the uniform, exponential, double exponential, Weibull, beta and the log-normal distributions. Upon completion of power computations, the

powers of the modified A-D and C-VM GOFTs will be compared to the unmodified A-D and C-VM GOFTs to determine which test is most powerful.

II. Literature Review.

2.1 Introduction

I have accomplished a literature review on the goodness-of-fit test techniques. Research studies have been investigated through the 1991 Current Index to Statistics (the most current available), the Defense Technical Information Center resources, and Air Force Institute of Technology theses. My discussion on hypothesis testing, test statistics, critical values, the bootstrap technique, plotting positions, power studies, and Monte Carlo simulation is important because these things are crucial to the conclusion of this thesis.

2.2 Goodness-of-fit Tests

Statistics is much more art than science. While scientists and engineers often devise and continually use statistical methods in their work, statistics is still a subjective art (17). The field of Statistics has many useful branches and procedures within it. As mentioned earlier, one of these areas has been in verifying a distribution for a set of observations. Over the years many different test procedures have been developed to accomplish this and they have come to be known as goodness- of-fit tests (7). The data are assumed to be samples from the hypothesized distribution unless there is sufficient evidence to disprove the assumption (1:72).

Similar to all other statistical work, there exist potential problems with the goodness-of-fit tests. One possible problem is that, even though one test shows there is insufficient evidence to disprove that the sample came from a certain distribution, there may be other distributions which fit the data better if only they are tested. Or it may not be an entirely different distribution, but the same distribution with different parameters. As such, it is important to remember that the goodness-of-fit test only rejects the assumed distribution when there is sufficient evidence that the distribution is incorrect; it does not mean that the selected distribution is the best

one (1:73). In order to perform a goodness-of-fit test, it is important to have an understanding of how a hypothesis test works.

2.3 Hypothesis Testing.

Hypothesis testing is the process of inferring from a sample whether to accept a certain statement about the population from which the sample was taken(5:76). Any statistical test of a hypothesis works in this way (16:429):

1. The null hypothesis, H_0, is stated in terms of the population.
2. The alternative hypothesis, H_a, is stated in terms of the population.
3. A test statistic is chosen.
4. A rule is made, in terms of possible values of the test statistic, for deciding when to accept the null hypothesis and when to reject it - the rejection region.
5. Based on a random sample from the population, the test statistic is evaluated, and a decision is made to accept or reject the null hypothesis.

The test statistic is a function of the sample data upon which the statistical decision will be based. If for a particular sample the computed value of the test statistic falls in the rejection region, the null hypothesis H_0 is rejected and the alternative hypothesis H_a is accepted. If the value of the test statistic does not fall into the rejection region, then H_0 is accepted (16:429)(5:77–78). Because statistics is not an exact science, it is not surprising that there are two ways of making a mistake in a hypothesis test. They are, either rejecting a true null hypothesis, or accepting a false null hypothesis. These two error types, types I and II respectively, have associated with them certain probabilities of their being made, designated as α and β (5: 79).

The following example illustrates how a statistical test works.

A psychological study was conducted to compare the reaction times of men and women to a certain stimulus. Independent random samples of 50 men and 50 women were employed in the experiment. The results are shown in table 2.1. Do the data present sufficient evidence to suggest a difference between true mean reaction times for men and women? Use $\alpha = .05$.

Table 2.1. Hypothesis Test Example Data

Men	Women
$n_1 = 50$	$n_2 = 50$
$\overline{y_1} = 3.6$ seconds	$\overline{y_2} = 3.8$ seconds
$s_1^2 = .18$	$s_2^2 = .14$

Solution: Let μ_1 and μ_2 denote the true mean reaction times for men and women, respectively. Then if we wish to test the hypothesis that the means are equal, we will test $H_0 : (\mu_1 - \mu_2) = 0$ against $H_a : (\mu_1 - \mu_2) \neq 0$. Note that we use the two-sided alternative to detect either the case $\mu_1 < \mu_2$ or the reverse $\mu_1 > \mu_2$, in case H_0 is false.

The point estimator of $(\mu_1 - \mu_2)$ is $(\overline{Y_1} - \overline{Y_2})$ and it satisfies the assumptions of our large-sample test. Hence if we desire to test $H_0 : \mu_1 - \mu_2 = D_0$ (D_0 fixed) versus any alternative, the test statistic is given by

$$Z = \frac{(\overline{Y_1} - \overline{Y_2}) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{2.1}$$

where σ_1^2 and σ_2^2 are the respective population variances. In this application we desire a two-tailed test. Thus for $\alpha = .05$ we reject H_0 for $|z| > z_{\frac{\alpha}{2}} = z_{.025} = 1.96$

For large samples (say, $n > 30$) the sample variances provide good estimates of their corresponding population variances. Substituting these values, along with $\overline{y_1}, \overline{y_2}, n_1, n_2$, and $D_0 = 0$, into the formula for the test statistic, we have $z = \frac{\overline{y_1} - \overline{y_2} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx \frac{.36 - .38}{\sqrt{\frac{.18}{50} + \frac{.14}{50}}} = -2.5$

This value is less than $-z_{\frac{\alpha}{2}} = -1.96$ and therefore falls in the rejection region. Hence we reject the hypothesis of no difference in mean reaction times for men and women (16:439–440).

There are many different types of hypothesis tests, one of the more common being of the type of the above example. This review is concerned with a more specific type of hypothesis test - the goodness-of-fit test. The goodness-of-fit tests require a

slightly different test statistic from the Z in the example above, which leads us to the next section - Test Statistics.

2.4 Test Statistics

Two major classes of test statistics used for goodness-of-fit tests are the Chi-square and the Empirical Distribution Function statistics.

2.4.1 Chi-square Test Statistics. The first goodness-of-fit test to be developed was the Chi-square test created by Karl Pearson in 1900. After working with data analysis for many years, he gave up on the notion that biological populations were normally distributed, he created other distributions, and then created his Chi-square test to determine which of these distributions a sample came from. Modern developments have increased the flexibility of the Chi-square tests, especially when unknown parameters must be estimated in the hypothesized family (7:63). The motivating force of the Chi-square test is that it compares the observed frequencies with expected frequencies of the hypothesized cumulative distribution function. This test is restricted to large samples, approximately 25 or greater (16:310). Even with the development of other goodness-of-fit tests, the advantages of the Chi-square test are not diminished; it is an economical first cut at trying to figure the samples distribution, it can be applied to discrete populations, and it can be modified when parameter values are unknown by reducing the number of degrees of freedom (15:68).

2.4.2 Empirical Distribution Function (EDF) Statistics. In addition to the basic Chi-square, there is the general class of statistics called Empirical Distribution Function statistics which are also used for the goodness-of-fit test. The empirical distribution function is a step function, taking a value between 0 and 1 for each member of the ordered sample, to estimate the cumulative distribution function. EDF statistics are just like their Chi-square counterparts in that they measure the discrepancy between the EDF and the hypothesized cumulative distribution function

(7: 97). For a given random sample of size n, let $X_{(1)}, ..., X_{(n)}$ be the ordered statistics. Suppose further that the distribution of X is F(x). Then, the empirical distribution function is $F_n(x)$ defined by

$$F_n(x) = \frac{number\ of\ observations \leq x}{n} \tag{2.2}$$

where

x ranges between $-\infty$ and ∞

More precisely, the definition becomes (7:98–99):

$$F_n(x) = 0, \quad x < X_{(1)} \tag{2.3}$$

$$F_n(x) = \frac{i}{n}, \quad X_{(i)} \leq x \leq X_{(i+1)}, \quad i = 1, ..., n-1 \tag{2.4}$$

where i is the rank of X_i amongst all the ordered observations.

$$F_n(x) = 1, \quad X_{(n)} \leq x \tag{2.5}$$

F_n(x) is a step function that takes a step up of height 1/n as each ordered sample observation is reached. One type of EDF statistic is the Kolmogorov-Smirnov statistic.

2.4.3 The Kolmogorov-Smirnov (K-S) Statistic. The K-S statistic is an EDF statistic in wide use. It is the largest vertical distance between the completely specified hypothesized cumulative distribution function (CDF) and the observed empirical distribution of the sample data(7:204). When it is necessary to completely specify all the parameters of the distribution, then use of the K-S statistic is appropriate, but only if the entire specification can be made prior to sampling (3:296).

Work done by D'Agostino and Stephens has shown that EDF statistics are much more powerful than Pearson Chi- square statistics (7:110). Another EDF statistic mentioned in the previous chapter is the C-VM statistic.

2.4.4 The Cramer-von Mises Statistic. The C-VM statistic is based on the squared summation of the difference between the EDF and the distribution being tested. This test statistic is defined as (7:101):

$$C - VM = \frac{1}{12n} + \sum_{i=1}^{n}(U_i - \frac{2i-1}{2n})^2 \qquad (2.6)$$

where n = sample size,

U_i = F($X_{(i)}$) = CDF for the distribution of interest,

i is the rank of X_i amongst all the ordered observations,

and $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are the n observations in order.

The C-VM test statistic, like the K-S test statistic, is also a function of the vertical distance between the hypothesized distribution and the Empirical Distribution Functions. They differ in the fact that the C-VM test goes one step further by considering every difference between the curve formed by the CDF and the curve formed by the EDF. Conover believes that the C-VM test is more appealing than the K-S test because the C-VM statistic uses more of the sample data. (5:306).

The following is a numerical example of how a C-VM goodness-of-fit test is performed on a sample of data, n=8:

H_0: The sample data in table 2.2 came from a normal distribution with $\mu = 4.3$ and $\sigma = 3.039$.
H_a: The sample came from some other population.
Decision Rule: If the C-VM statistic $\geq$ .1182 reject the null hypothesis, otherwise accept the null hypothesis. (note: using α = .05 and n=8 the C-VM critical value from the table on page 4-3 is .1182)

The cumulative distribution function for the normal distribution being:

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad (2.7)$$

And remember that $C - VM = \frac{1}{12n} + \sum_{i=1}^{n}(U_i - \frac{2i-1}{2n})^2$

Table 2.2. C-VM GOFT Example Data

j	x	F(x)=U_j	$A_j = \frac{2j-1}{2n}$	$(U_j - A_j)^2$
1	0.2	.1038	.0625	.00171
2	1.6	.2033	.1875	.00025
3	2.1	.2483	.3125	.00412
4	3.0	.3446	.4375	.00863
5	4.8	.5596	.5625	.00001
6	5.0	.5871	.6875	.01008
7	8.1	.8790	.8125	.00442
8	9.6	.9484	.9375	.00012

$\sum_{j=1}^{8}(U_j - A_j)^2 = .02934$

$C - VM = \frac{1}{12*8} + .02934 = .01042 + .02934 = .03976$

Because .03976 $\ngtr$ .1182, we accept the null hypothesis that the sample does indeed come from a normal distribution with $\mu = 4.3$ and $\sigma = 3.039$ (22:17–18).

Another EDF statistic mentioned in the previous chapter is the A-D statistic.

2.4.5 The Anderson-Darling Statistic. Anderson and Darling created a new statistic by incorporating a weight function into the K-S and C-VM statistics. The A-D test statistic is defined as (7:100–101):

$$A - D = -n - 2\sum_{i=1}^{n} \frac{2i-1}{2n}(\log_e(U_i) + log_e(1 - U_{n-i+1})) \quad (2.8)$$

where n = sample size,

U_i = F($X_{(i)}$) = CDF for the distribution of interest,

i is the rank of X_i amongst all the ordered observations,

and $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ are n observations in order.

The A-D statistic is based on a weighted average of the squared discrepancies between the curve formed by the CDF and the curve formed by the EDF.

The following is a numerical example of how an A-D goodness-of-fit test is performed on a sample of data, n=8:

H_0: The sample data in table 2.3 came from a normal distribution with $\mu = 4.3$ and $\sigma = 3.039$.

H_a: The sample came from some other population.

Decision Rule: If the A-D statistic $\geq .6549$ reject the null hypothesis, otherwise accept the null hypothesis. (note: using $\alpha = .05$ and n=8 the A-D critical value from the table on page 4-2 is .6549)

Remember that $A-D = -n - 2\sum_{i=1}^{n} \frac{2i-1}{2n}(\log_e(U_i) + log_e(1-U_{n-i+1}))$ and $F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$. Also, $A_i = \ln(U_i)$ and $B_i = \ln(1 - U_{n-i+1})$

Table 2.3. A-D GOFT Example Data

j	x	F(x)=U_j	U_{n-j+1}	A_j	B_j	$(2j-1)(A_j+B_j)$
1	0.2	.1038	.9484	-2.265	-2.964	-5.229
2	1.6	.2033	.8790	-1.593	-2.112	-11.115
3	2.1	.2483	.5871	-1.393	-.885	-11.390
4	3.0	.3446	.5596	-1.065	-.820	-13.195
5	4.8	.5596	.3446	-.581	-.423	-9.036
6	5.0	.5871	.2483	-.533	-.285	-8.998
7	8.1	.8790	.2033	-.129	-.227	-4.628
8	9.6	.9484	.1038	-.053	-.110	-2.445

$\sum_{j=1}^{8}(2j-1)(A_j + B_j) = .02934$

$A - D = -8 - (\frac{1}{8})(-66.036) = -8 + 8.2545 = .2545$ Because $.2545 \not\geq .6549$, we accept the null hypothesis that the sample does indeed come from a normal distribution with $\mu = 4.3$ and $\sigma = 3.039$ (22:16–17).

The critical values used in the two tests above were taken from critical value tables, which leads to the next part of this review, critical values.

2.5 Critical Values.

Previous critical values have been generated by taking random samples from the test distribution, calculating the test statistic and then ordering the statistics

from smallest to largest. The best way to explain the nature of a critical value is with a word picture: if 100 independent random samples of size n are generated, and from them 100 goodness-of-fit statistics are calculated and ordered from smallest to largest, then the critical value for $1 - \alpha = 1 - .05 = .95$ is the 95^{th} largest order statistic. This reasoning works well when there is a one-to-one correspondence between the α-level and the order statistic, but what if for the example with 100 statistics a critical value was needed at $\alpha = .0005$? There would be a problem because in the example, discrete values are being used to determine critical values for what is actually a continuous distribution. To get around this conflict, all that must be done is an interpolation of the data, and for endpoints, an extrapolation. This is simply a method for representing these order statistics on a continuous spectrum. This is done by plotting the values of the order statistics and representing the spaces between them as piecewise linear functions. To find a critical value, all that is necessary, graphically, is to find the value on the x-axis that corresponds with the selected $1 - \alpha$ value of interest on the y-axis (8).

First, the n ordered test statistics, $x_1, x_2, ..., x_n$, are plotted against their $1-\alpha = y_1, y_2, ..., y_n$ which are calculated with the median rank formula of the form: $y_i = \frac{i-.3}{n+.4}$, i=the rank of the ordered statistic and n=sample size. Letting the ordered statistics be represented by the horizontal axis and letting the y_i be represented by the vertical axis and not coincidentally be scaled between zero and one, the statistics will be represented on a continuous function. After this is accomplished, the endpoints are found by linear extrapolation. The first point on the horizontal axis, x_0, is computed by linearly extrapolating from the second and third points, subject to a non-negativity restriction. Extrapolation is performed by using the standard linear slope-intercept formula $y = m * x + b$ to compute the endpoints x_0 and x_{n+1}. To find the first endpoint on the horizontal axis, the slope is calculated by:

$$m = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \tag{2.9}$$

and the intercept is found by:

$$b = y_i - m * x_i \tag{2.10}$$

Then the lower endpoint x_0 is found by:

$$x_0 = \frac{y_0 - b}{m} = \frac{0 - b}{m} = \frac{-b}{m} \tag{2.11}$$

Using the non-negativity rule,

$$x_0 = max(0, \frac{-b}{m}) \tag{2.12}$$

After both endpoints are calculated, n test statistic values and their two extrapolated endpoints are plotted on the x - y axis. Then, the critical values corresponding to the desired percentiles are found by linearly interpolating between the points (x_i, y_i) and (x_{i+1}, y_{i+1}) using the formulas above; where C_p is the critical value for the desired percentile (20).

> To find a critical value, all that is necessary, graphically, is to find $1 - \alpha$ on the vertical axis and extend along the line, $Y_{(i)} = 1 - \alpha$, to intercept the plotted function. The value of the horizontal component is the critical value of the statistic at significance level α.
>
> Finding the critical value with a computer requires finding the largest $Y_{(i)}$ that is less than $1-\alpha$. Suppose that $Y_{(k)}$ is the k^{th} largest rank. Then, the standard linear slope-intercept formula ($y = mx + b$) is used to find the critical value. The change in y can be found using $Y_{(k)}$ and $Y_{(k+1)}$. Similarly, the change in x can be found using $X_{(k)}$ and $X_{(k+1)}$. After finding the constant, b, at $(X_{(k)}, Y_{(k)})$, one can then let y equal $1 - \alpha$ in order to find x, the critical value.
>
> Suppose ten samples are taken as shown in table 2.4:
>
> In figure 2.1, the statistics are plotted versus their median step ranks. And from table 2.4, $Y_{(1)} = .05$ $Y_{(2)} = .15$ $X_{(1)} = .22$ $X_{(2)} = .41$ Using the equation, $y = mx + b$, $m = slope = \frac{Y_{(2)} - Y_{(1)}}{X_{(2)} - X_{(1)}} = \frac{.163 - .067}{.41 - .22} = .505$ and $b = Y_{(1)} - mX_{(1)}$ $= .067 - (.505)(.22) = -.0441$ and $x = \frac{0-b}{m} = \frac{.0441}{.505} =$

Table 2.4. Critical-value Example Data

i	Median Rank $y_i = \frac{i-.3}{n+.4}$	Statistics x_i
1	.067	.22
2	.163	.41
3	.26	.42
4	.356	.67
5	.452	.98
6	.548	1.02
7	.644	1.03
8	.74	1.08
9	.837	1.12
10	.933	1.13

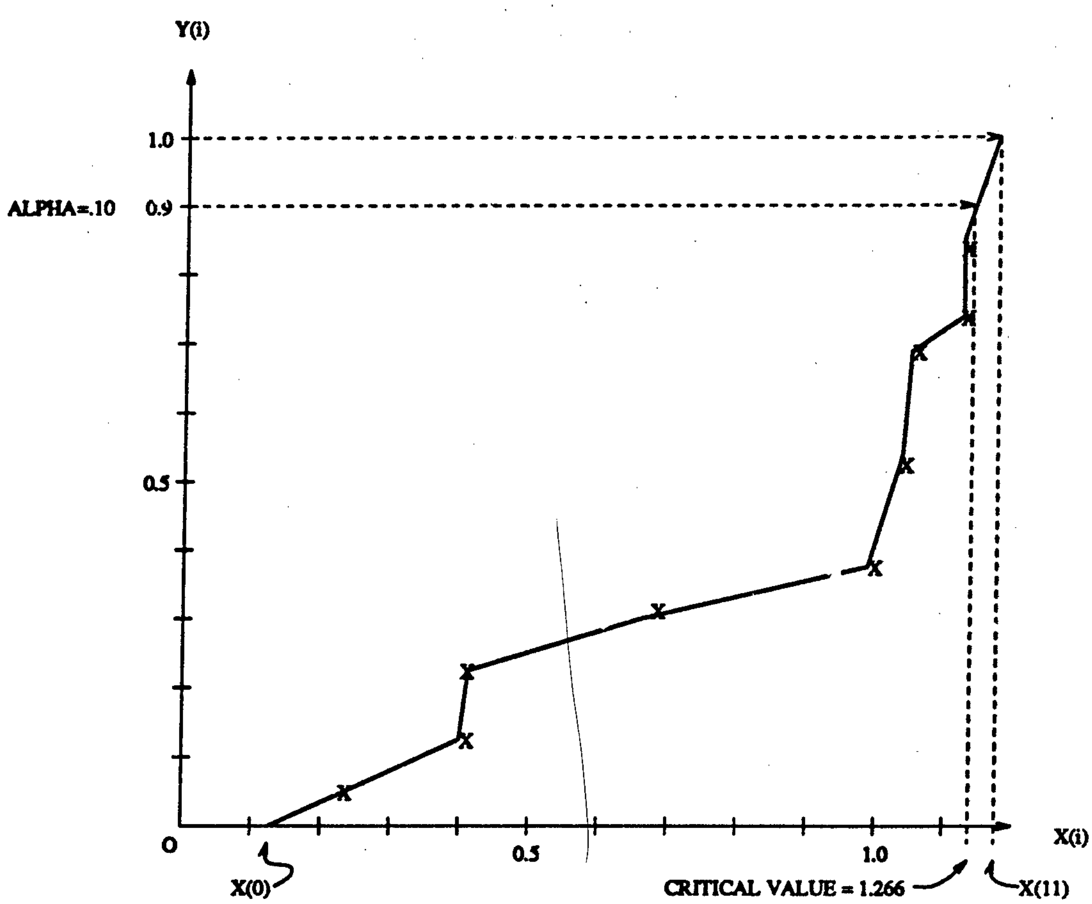

Figure 2.1. Critical Value Example

.0873 Since $x = .0873 \geq 0$, $X_{(0)} = x$. Again, if x had been less than zero, $X_{(0)}$ would have been set to zero. Extrapolation for $X_{(11)}$ is performed the same way. $Y_{(10)}$, $Y_{(9)}$, $X_{(10)}$, and $X_{(9)}$ are used to find the slope. The constant, b, is calculated at either $(X_{(10)}, Y_{(10)})$ or at $(X_{(9)}, Y_{(9)})$. Then $X_{(11)} = \frac{1-b}{m}$, where m is the slope. Now that the function is continuous (by extrapolation) on the interval (0,1), the critical values can be found. At $\alpha = .10$, previous studies would have picked 1.12, or the 9^{th} largest statistic as the critical value. Using the bootstrap method, the value is 1.1266 (if the given median step rank is used). To get the critical value using the bootstrap technique, the largest $Y_{(i)}$ less than or equal to .90 is found. In this case, this is $Y_{(9)} = .85$. Therefore, k=9 and $k+1 = 10$. Then, $m = \frac{Y_{(10)}-Y_{(9)}}{X_{(10)}-X_{(9)}} = \frac{.933-.837}{1.13-1.12} = \frac{.096}{.01} = 9.6$ $b = .837 - (9.6)(1.12) = .837 - 10.752 = -9.915$ and $critical\ value = \frac{.90-(-9.915)}{9.6} = 1.1266$ As one can see, the critical value will vary with statistics calculated from random samples. (22:19-23).

Typically, researchers are interested in getting the critical values for several α-levels such as $\alpha = .01, ..., .20$. Once these critical value tables have been generated and tabled for a particular distribution, an assessment of the usefulness of the table needs to be performed. This assessment has come to be known as a power study.

2.6 Power Studies.

The power of the goodness-of-fit test is the probability that the test will reject the null hypothesis when in fact it should reject the null hypothesis (16:473). The statistician wants his test to be as powerful as possible for identifying those estimates of distributions that are flat-out wrong. The following discussion presents some of the work done in efforts to create more powerful goodness-of-fit tests.

Green and Hegazy teamed-up to attempt to improve the standard EDF goodness-of-fit tests. They modified the A-D, C-VM, and the K-S goodness-of-fit tests where the parameters were estimated from the sample. They showed that the modified A-D and C-VM tests achieved higher power (9:204).

Porter modified the A-D, C-VM and the K-S goodness-of-fit tests for the three-parameter Pareto distribution. He concluded that the power of his modified tests

improved as the sample sizes increased. He also showed that A-D, C-VM and K-S tests were more powerful than the Chi-square, especially when the sample data were taken from the Weibull, beta, or normal distributions (20:7).

Ream modified the K-S, A-D, and C-VM tests for the normal distribution to test the technique of reflecting data points about the mean and created tables of critical values of the modified K-S, A-D, and C-VM statistics. He concluded that against symmetric distributions the power increased, whereas against non-symmetric distributions a significant loss of power occurred. He further recommended that the consequences of this fact be strongly considered before use of his new technique and statistics (22:62–63).

Bush modified the A-D and C-VM tests for the Weibull distribution. He presented a Monte Carlo method for obtaining the critical values of the modified A-D and C-VM goodness-of-fit tests for the three parameter Weibull distribution when the scale and location parameters were not specified. He also concluded that A-D and C-VM tests were not very powerful when the sample size was five. When the sample size increased, he showed that A- D and C-VM were more powerful than K-S tests (4:52).

Viviano modified the A-D and C-VM tests for the Gamma distribution. He showed that the A-D goodness-of-fit test was more powerful against the log-normal distribution than the C-VM test(24:47).

Kahya modified the A-D and C-VM tests by adjusting the plotting positions of the original statistics. He showed that the power of the A-D test increased for most all distributions and the power of the C-VM test increased for the uniform, Weibull, exponential, double exponential, and log-normal (13:6–1).

According to Stephens, critical value tables generated from one continuous distribution can be applied to GOFTs for all continuous distributions as long as the parameters are completely specified for the sample (7:102).

All work with GOFTs involves many tests. In fact, to do work of any significance, many tests must be performed on samples of data. And that leads us to the last topic of this chapter, Monte Carlo simulation.

2.7 Monte Carlo Simulation

In many cases a system or phenomenon under study is either too complex to be adequately defined by mathematical formulation or the mathematical formulation is extremely difficult, or even impossible, to solve by analytic techniques. When these difficulties occur, a solution can often be obtained by use of the technique mentioned earlier - simulation. Simulation is a procedure for imitating (not representing) a system or phenomenon for different values of the controllable variables. Much of the work with GOFTs use Monte Carlo simulation to create data that mimics many different populations: e.g. normal, log-normal, exponential, double exponential, uniform, and Weibull. It is one thing to have data that is representative of a distribution, e.g. ten samples from an exponential distribution with $\lambda = 10$, but it is quite another to have that sample be randomly generated again and again and again (17). The approach is to observe random variates, chosen so that they directly simulate the random processes of the original problem. Then based upon the multiple observations, the desired general solution is inferred from the behavior of the random numbers (10:2–4). The main weakness in Monte Carlo simulation is that the answers it produces are to some degree uncertain since they are inferred raw observational data consisting of raw numbers. This weakness must be accounted for because:

> Whenever one is inferring general laws on the basis of particular observations associated with them, the conclusions are uncertain inasmuch as the particular observations are only a more or less representative sample from the totality of all observations which might have been made. Good experimentation tries to ensure that the sample shall be more rather than less representative... [Monte Carlo answers] can nevertheless serve a useful purpose if we can manage to make the uncertainty fairly negligible,

that is to say to make it unlikely that the answers are wrong by very much (10:4–5).

Thus there is usually no cause for concern if the uncertainty is negligible for practical purposes. One way of reducing uncertainty is to base the Monte Carlo work on a larger number of observations. This is done because the law of large numbers states that, as the sample size increases, the differences between the sample mean and the population mean becomes smaller (1:176).

All work done in this area has not been done to satisfy one particular sample of data from say the Weibull distribution. But instead, the researchers were trying to establish general principles that applied, if not across all probability distribution families, then at least across one distribution family - regardless of sample size. To achieve this end, Monte Carlo simulation was key.

2.8 Conclusion.

Every single one of these research studies has been done in order to try to find more powerful tests for the different distributions. They all modified the test techniques or used a new technique for only one purpose; to obtain a "better" test so that the analysts are better able to determine how well a set of data fits some distribution. However, there are still many different avenues that can be pursued to find even more powerful tests for determining goodness-of-fit. Research must be done in this area.

III. Methodology

3.1 Introduction

This chapter describes the procedures used to accomplish the research objectives of this thesis. The discussion covers the modification of the A-D and C-VM statistics, calculation of the critical values, and finally, calculation of the power tables.

3.2 Discussion

The research effort will consist of the following parts:

1. The A-D and C-VM critical value tables will be derived using the known A-D and C-VM statistic formulas.
2. The A-D and C-VM critical value tables will be generated using the modified A-D and C-VM statistic formulas.
3. The powers of the modified A-D and C-VM statistics will be determined by Monte Carlo simulation testing against alternative distributions. These alternative distributions will be the uniform, exponential, double exponential, Weibull, beta, and log-normal distributions.
4. Analysis will be conducted on the powers of each alternative.

3.2.1 Modified Test Statistics. As stated earlier, the work of this thesis is to determine which A-D and C-VM goodness-of-fit statistics are most powerful. In addition to examining the power of the unmodified A-D and C-VM statistics, this thesis will examine five modifications of each of the A-D and C-VM statistics (a total of ten new statistics). This research will substitute in place of the $\frac{2i-1}{2n}$ portion of the statistics this median rank quotient $P_i = \frac{i-c}{n-2c+1}$, i=rank of the ordered statistic, and n=sample size. The research will use c = 0, 0.3, 0.3175, 0.375 and 1 to generate

these five different median rank plotting positions for development of the modified A-D and C-VM statistics:

$$P_i = \frac{i}{n+1} \tag{3.1}$$

$$P_i = \frac{i-0.3}{n+0.4} \tag{3.2}$$

$$P_i = \frac{i-0.3175}{n+0.365} \tag{3.3}$$

$$P_i = \frac{i-0.375}{n+0.25} \tag{3.4}$$

$$P_i = \frac{i-1}{n-1} \tag{3.5}$$

Using these plotting positions, the modified A-D and C-VM statistics are:

$$AD_{MEDIAN,C=0} = -n - 2\sum_{i=1}^{n} \frac{i}{n+1}(\log_e(U_i) + log_e(1 - U_{n-i+1})) \tag{3.6}$$

$$AD_{MEDIAN,C=0.3} = -n - 2\sum_{i=1}^{n} \frac{i-0.3}{n+0.4}(\log_e(U_i) + log_e(1 - U_{n-i+1})) \tag{3.7}$$

$$AD_{MEDIAN,C=0.3175} = -n - 2\sum_{i=1}^{n} \frac{i-0.3175}{n+0.365}(\log_e(U_i) + log_e(1 - U_{n-i+1})) \tag{3.8}$$

$$AD_{MEDIAN,C=0.375} = -n - 2\sum_{i=1}^{n} \frac{i-0.375}{n+0.25}(\log_e(U_i) + log_e(1-U_{n-i+1})) \quad (3.9)$$

$$AD_{MEDIAN,C=1} = -n - 2\sum_{i=1}^{n} \frac{i-1}{n-1}(\log_e(U_i) + log_e(1-U_{n-i+1})) \quad (3.10)$$

$$CVM_{MEDIAN,C=0} = \frac{1}{12n} + \frac{1}{n}\sum_{i=1}^{n}(U_i - \frac{i}{n+1})^2 \quad (3.11)$$

$$CVM_{MEDIAN,C=0.3} = \frac{1}{12n} + \frac{1}{n}\sum_{i=1}^{n}(U_i - \frac{i-0.3}{n+0.4})^2 \quad (3.12)$$

$$CVM_{MEDIAN,C=0.3175} = \frac{1}{12n} + \frac{1}{n}\sum_{i=1}^{n}(U_i - \frac{i-0.3175}{n+0.365})^2 \quad (3.13)$$

$$CVM_{MEDIAN,C=0.375} = \frac{1}{12n} + \frac{1}{n}\sum_{i=1}^{n}(U_i - \frac{i-0.375}{n+0.25})^2 \quad (3.14)$$

$$CVM_{MEDIAN,C=1} = \frac{1}{12n} + \frac{1}{n}\sum_{i=1}^{n}(U_i - \frac{i-1}{n-1})^2 \quad (3.15)$$

3.2.2 Calculation of the Critical Values. The critical value tables will be generated using Monte Carlo simulation of random deviates. The following steps describe the procedure for generating the critical value table for the unmodified A-D when the sample size is four.

1. Using the International Mathematical and Statistics Library (IMSL) subroutine RNNOR, the random variates will be generated from the normal distribution with a mean of $\mu = 0$ and a standard deviation of $\sigma = 1$.

2. The random variates will be standardized by using the following transformation:

$$Z_i = \frac{x_i - \overline{x}}{s} \qquad (3.16)$$

where $\overline{x} = \frac{\sum_{i=1}^{n}}{n}$, the sample mean, and $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$, the unbiased sample standard deviation.

These standardized data will be used in steps three through six.

3. The hypothesized cumulative distribution function for the normal distribution, U_i, will then be calculated for $i = 1, 2, ..., n$.

$$U_i = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2s^2}(x-\overline{x})^2} \qquad (3.17)$$

4. The A-D statistic will be calculated.
5. Steps one through four will be repeated 9,999 more times to generate a total of 10,000 independent A-D statistics.
6. These 10,000 statistics will be ordered from smallest to largest. Using the bootstrap technique, the $80^{th}, 85^{th}, 90^{th}, 95^{th}$, and 99^{th} percentiles of the distributions of the A-D statistic will be determined by linear interpolations. These percentiles represent the .20, .15, .10, .05, and .01 levels of significance. Then these values will be entered into the A-D critical value table.

These six steps will be done again for $n = 5, ..., 50$ for the unmodified A-D statistic. Then the whole procedure will be repeated for the five modifications to the A-D statistic, the unmodified C-VM statistic, and the five modified C-VM statistics as well. The flowchart representation for generating the critical values is shown in figure 3.1.

3.2.3 Power of the Goodness-of-fit Tests. Basically, the power of a goodness-of-fit test is the probability that the test will reject the null hypothesis when in fact

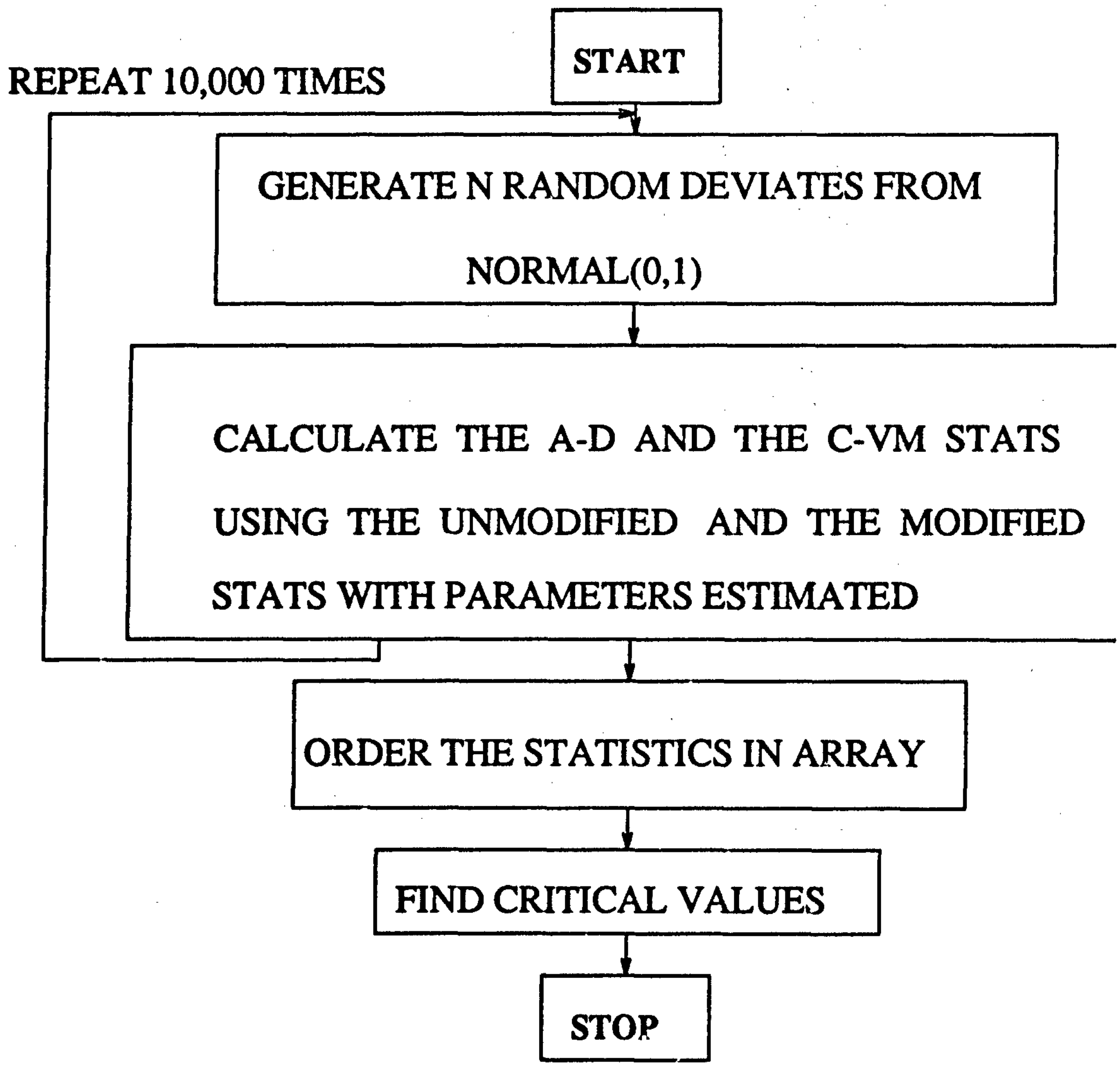

Figure 3.1. Generation of Critical Values

it should be rejected. The reason for conducting the power test is to help determine if the modified A-D and C-VM statistics result in more powerful goodness-of-fit tests than the unmodified A-D and C-VM statistics.

In order to run a power test, random deviates from non-Normal distributions of sample size n will be generated and $\bar{x}$ and s will be calculated. Then, the A-D and C-VM statistics will be calculated using the normal CDF. Because of the earlier work, the critical value tables are available for look-up of critical values by sample size, n, and by probability of a type I error, α. The calculated A-D and C-VM statistics will then be compared to their respective table value to see if the test passes or fails. This process will be repeated 10,000 times and the number of times each statistic value exceeds the respective critical value will be counted for each sample size of n. This total number will represent the number of the rejections of the null hypothesis. Ideally, this count would be 10,000, meaning none of the samples passed the test for coming from a normal distribution, but this will not be the case as the test is far from clairvoyant. This total rejection number will be divided by 10,000 to obtain the power

$$The\ power = \frac{Total\ number\ of\ rejections.}{10,000} \tag{3.18}$$

The power study is performed at n=10, n=20, n=30, n=40, and n=50.After getting all the powers at each α-level, final conclusions will be derived for each alternative distribution.

To reiterate, the power study will be conducted using the following test and the steps in figure 3.2:

H_o : Sample deviates follow a normal distribution.

H_a : They follow some other distribution.

Decision Rule: If the A-D statistic $\geq$ *critical value* reject the null hypothesis, otherwise accept the null hypothesis

3.3 Conclusion

This chapter outlined the specific steps that will be taken to generate the tests and data needed to accomplish this thesis. The next chapter will show the tables generated by pursuing the methods of this chapter.

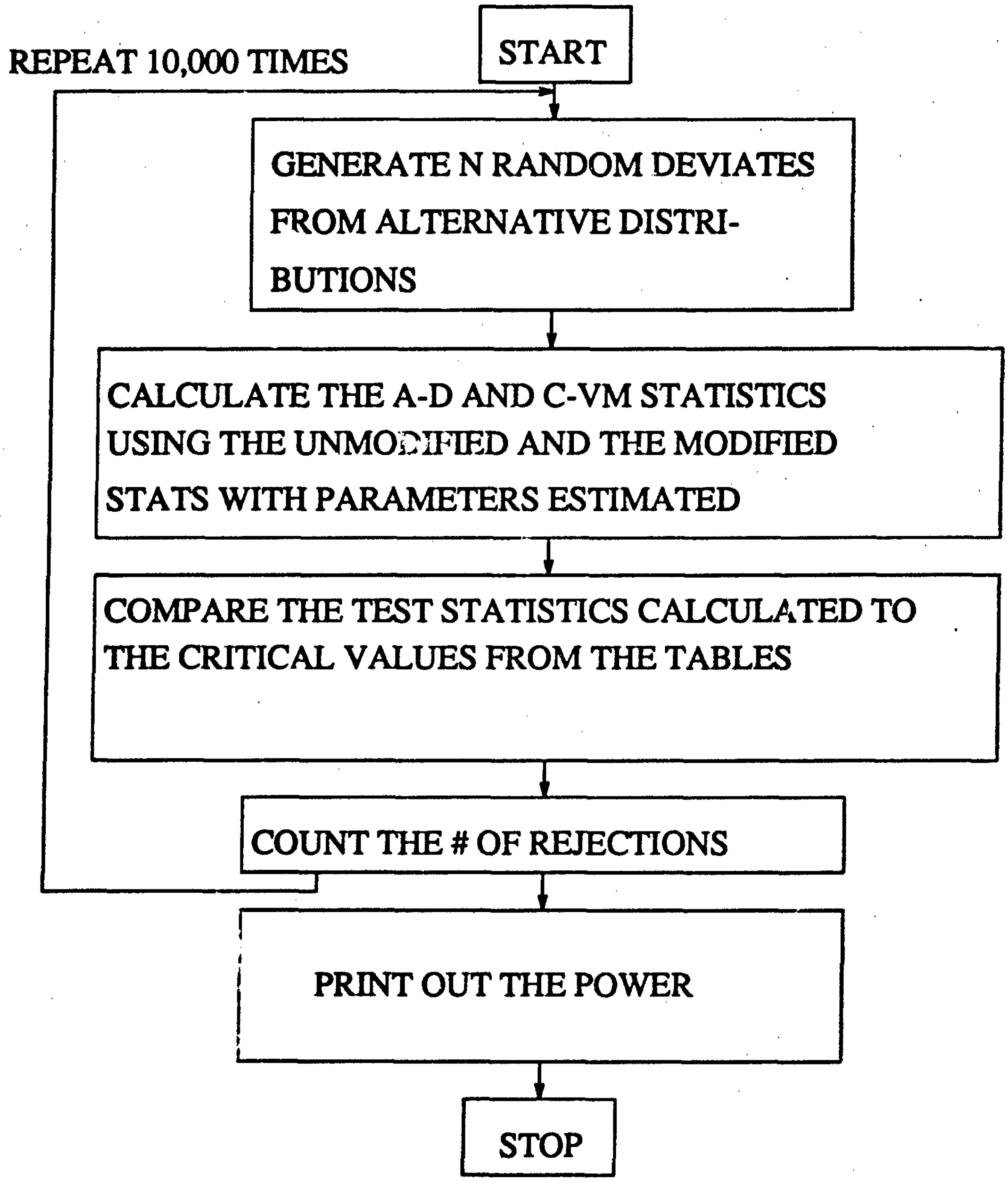

Figure 3.2. Generation of Power tables.

IV. Results

Following the steps outlined in the previous chapter, this chapter presents the critical value tables and the power study tables.

4.1 Critical Value Tables

To use these tables, select the α level from across the top row, and then go down the table to the appropriate sample size to get the correct critical value for the test.

Table 4.1. Critical Values for the ANDERSON-DARLING Test with Unmodified Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.4033	0.4399	0.4393	0.5669	0.6954
5	0.4286	0.4701	0.5222	0.5998	0.7971
6	0.44[illegible]2	0.4846	0.5386	0.6287	0.8503
7	0.4493	0.4947	0.5532	0.6589	0.8806
8	0.4599	0.5041	0.5608	0.6549	0.8942
9	0.4636	0.5145	0.5722	0.6847	0.9457
10	0.4766	0.5234	0.5831	0.6924	0.9265
11	0.4781	0.5226	0.5854	0.6951	0.9777
12	0.4766	0.5228	0.5936	0.7142	0.9835
13	0.4774	0.5254	0.5881	0.7028	0.9761
14	0.4835	0.5297	0.5976	0.7214	0.9829
15	0.4869	0.5370	0.6016	0.7125	0.9996
16	0.4820	0.5282	0.5985	0.7222	0.9550
17	0.4928	0.5420	0.6102	0.7191	0.9604
18	0.4901	0.5379	0.6061	0.7270	0.9910
19	0.4882	0.5341	0.5988	0.7245	0.9817
20	0.4937	0.5390	0.6055	0.7171	0.9956
21	0.4977	0.5436	0.6162	0.7299	0.9951
22	0.4900	0.5356	0.6031	0.7265	1.0012
23	0.4926	0.5506	0.6152	0.7214	1.0034
24	0.4920	0.5418	0.6064	0.7190	0.9831
25	0.4924	0.5410	0.6160	0.7296	1.0153
26	0.4934	0.5416	0.6063	0.7142	0.9966
27	0.4964	0.5491	0.6158	0.7412	1.0184
28	0.4965	0.5480	0.6188	0.7330	1.0311
29	0.4993	0.5474	0.6130	0.7322	1.0086
30	0.4960	0.5464	0.6164	0.7313	1.0083
31	0.4980	0.5530	0.6217	0.7458	1.0252
32	0.4998	0.5529	0.6230	0.7424	1.0117
33	0.4981	0.5484	0.6169	0.7342	1.0003
34	0.4953	0.5464	0.6146	0.7312	1.0104
35	0.5032	0.5493	0.6203	0.7371	1.0199
36	0.4967	0.5482	0.6166	0.7369	0.9[illegible]69
37	0.4949	0.5458	0.6102	0.7256	1.0008
38	0.5010	0.5473	0.6151	0.7348	1.0103
39	0.5033	0.5502	0.6191	0.7375	1.0070
40	0.5033	0.5513	0.6252	0.7445	1.0143
41	0.4977	0.5467	0.6183	0.7283	1.0130
42	0.4995	0.5505	0.6171	0.7308	1.0235
43	0.4981	0.5500	0.6223	0.7484	1.0154
44	0.5010	0.5491	0.6185	0.7459	1.0417
45	0.5033	0.5512	0.6196	0.7470	1.0369
46	0.5008	0.5480	0.6206	0.7325	1.0029
47	0.4977	0.5477	0.6223	0.7384	1.0116
48	0.5072	0.5554	0.6218	0.7370	1.0252
49	0.4958	0.5465	0.6193	0.7457	1.0221
50	0.5010	0.5539	0.6251	0.7484	1.0381

Table 4.2. Critical Values for the CRAMER-VON MISES Test with Unmodified Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.0702	0.0771	0.0871	0.1057	0.1348
5	0.0738	0.0813	0.0916	0.1083	0.1526
6	0.0758	0.0836	0.0954	0.1154	0.1608
7	0.0764	0.0851	0.0967	0.1164	0.1632
8	0.0775	0.0855	0.0970	0.1182	0.1641
9	0.0767	0.0850	0.0975	0.1162	0.1637
10	0.0775	0.0855	0.0979	0.1192	0.1712
11	0.0788	0.0879	0.1004	0.1202	0.1714
12	0.0796	0.0889	0.1000	0.1223	0.1690
13	0.0793	0.0883	0.1000	0.1213	0.1662
14	0.0791	0.0886	0.1008	0.1212	0.1744
15	0.0796	0.0887	0.1012	0.1216	0.1751
16	0.0792	0.0879	0.1011	0.1252	0.1735
17	0.0796	0.0876	0.1000	0.1205	0.1750
18	0.0796	0.0887	0.1013	0.1232	0.1775
19	0.0796	0.0886	0.1012	0.1258	0.1829
20	0.0789	0.0877	0.1012	0.1217	0.1711
21	0.0801	0.0893	0.1024	0.1247	0.1725
22	0.0803	0.0892	0.1009	0.1230	0.1696
23	0.0796	0.0887	0.1024	0.1238	0.1754
24	0.0801	0.0888	0.1020	0.1244	0.1715
25	0.0792	0.0880	0.1010	0.1231	0.1736
26	0.0795	0.0890	0.1025	0.1233	0.1695
27	0.0802	0.0896	0.1020	0.1233	0.1732
28	0.0804	0.0906	0.1038	0.1256	0.1752
29	0.0804	0.0890	0.1020	0.1226	0.1787
30	0.0803	0.0890	0.1019	0.1258	0.1775
31	0.0800	0.0890	0.1010	0.1230	0.1691
32	0.0802	0.0884	0.1012	0.1240	0.1781
33	0.0804	0.0893	0.1021	0.1255	0.1726
34	0.0806	0.0899	0.1030	0.1238	0.1737
35	0.0801	0.0897	0.1022	0.1244	0.1791
36	0.0801	0.0894	0.1015	0.1249	0.1737
37	0.0801	0.0899	0.1029	0.1253	0.1732
38	0.0803	0.0898	0.1025	0.1235	0.1728
39	0.0808	0.0897	0.1023	0.1234	0.1715
40	0.0811	0.0906	0.1028	0.1249	0.1796
41	0.0806	0.0902	0.1032	0.1236	0.1755
42	0.0803	0.0891	0.1015	0.1238	0.178[illegible]
43	0.0794	0.0886	0.1013	0.1233	0.1775
44	0.0803	0.0900	0.1035	0.1262	0.1779
45	0.0808	0.0901	0.1017	0.1240	0.1742
46	0.0804	0.0895	0.1028	0.1232	0.1769
47	0.0807	0.0898	0.1030	0.1257	0.1803
48	0.0812	0.0899	0.1023	0.1236	0.1783
49	0.0814	0.0905	0.1030	0.1244	0.1794
50	0.0803	0.0895	0.1018	0.1265	0.1754

Table 4.3. Critical Values for the ANDERSON-DARLING Test with C = 0.3175 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.6555	0.6892	0.7340	0.8047	0.9224
5	0.6999	0.7384	0.7878	0.8589	1.0419
6	0.7288	0.7680	0.8184	0.9033	1.1101
7	0.7454	0.7887	0.8439	0.9441	1.1530
8	0.7640	0.8054	0.8592	0.9502	1.1770
9	0.7783	0.8230	0.8776	0.9866	1.2350
10	0.7919	0.8373	0.8930	0.9991	1.2245
11	0.7976	0.8406	0.9012	1.0067	1.2779
12	0.7992	0.8447	0.9129	1.0291	1.2911
13	0.8041	0.8497	0.9104	1.0214	1.2873
14	0.8116	0.8571	0.9226	1.0446	1.2946
15	0.8175	0.8654	0.9299	1.0372	1.3176
16	0.8159	0.8598	0.9294	1.0494	1.2760
17	0.8270	0.8762	0.9421	1.0487	1.2876
18	0.8270	0.8742	0.9402	1.0566	1.3145
19	0.8259	0.8708	0.9343	1.0574	1.3077
20	0.8328	0.8771	0.9425	1.0511	1.3229
21	0.8383	0.8834	0.9534	1.0646	1.3256
22	0.8313	0.8756	0.9429	1.0641	1.3330
23	0.8354	0.8915	0.9557	1.0598	1.3354
24	0.8357	0.8844	0.9473	1.0593	1.3200
25	0.8366	0.8842	0.9587	1.0700	1.3522
26	0.8387	0.8859	0.9499	1.0549	1.3317
27	0.8416	0.8939	0.9604	1.0840	1.3539
28	0.8424	0.8935	0.9627	1.0753	1.3701
29	0.8461	0.8940	0.9576	1.0761	1.3477
30	0.8427	0.8932	0.9627	1.0756	1.3483
31	0.8463	0.9005	0.9683	1.0905	1.3654
32	0.8481	0.9014	0.9703	1.0875	1.3525
33	0.8470	0.8968	0.9656	1.0810	1.3433
34	0.8450	0.8946	0.9621	1.0781	1.3546
35	0.8530	0.8990	0.9694	1.0837	1.3626
36	0.8478	0.8972	0.9659	1.0851	1.3394
37	0.8458	0.8955	0.9598	1.0741	1.3434
38	0.8520	0.8986	0.9653	1.0843	1.3553
39	0.8548	0.9012	0.9690	1.0857	1.3504
40	0.8552	0.9028	0.9757	1.0933	1.3600
41	0.8504	0.8987	0.9691	1.0784	1.3591
42	0.8524	0.9023	0.9682	1.0816	1.3715
43	0.8509	0.9021	0.9737	1.0983	1.3627
44	0.8543	0.9020	0.9704	1.0974	1.3909
45	0.8561	0.9043	0.9713	1.0987	1.3864
46	0.8537	0.9007	0.9726	1.0831	1.3506
47	0.8518	0.9008	0.9749	1.0896	1.3612
48	0.8615	0.9088	0.9748	1.0893	1.3759
49	0.8503	0.9003	0.9726	1.0973	1.3706
50	0.8555	0.9078	0.9779	1.1012	1.3852

Table 4.4. Critical Values for the CRAMER-VON MISES Test with C = 0.3175 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.0694	0.0758	0.0842	0.0997	0.1262
5	0.0721	0.0793	0.0881	0.1034	0.1411
6	0.0740	0.0814	0.0916	0.1103	0.1505
7	0.0749	0.0825	0.0938	0.1116	0.1528
8	0.0755	0.0831	0.0941	0.1135	0.1552
9	0.0753	0.0830	0.0946	0.1112	0.1558
10	0.0759	0.0836	0.0947	0.1148	0.1637
11	0.0774	0.0857	0.0974	0.1163	0.1640
12	0.0783	0.0871	0.0977	0.1189	0.1627
13	0.0780	0.0865	0.0984	0.1173	0.1591
14	0.0778	0.0872	0.0991	0.1182	0.1684
15	0.0784	0.0870	0.0990	0.1182	0.1687
16	0.0785	0.0865	0.0989	0.1217	0.1685
17	0.0786	0.0866	0.0978	0.1180	0.1686
18	0.0789	0.0876	0.0997	0.1209	0.1720
19	0.0788	0.0874	0.0995	0.1223	0.1774
20	0.0782	0.0865	0.0998	0.1203	0.1681
21	0.0793	0.0884	0.1009	0.1223	0.1662
22	0.0796	0.0878	0.0997	0.1204	0.1653
23	0.0789	0.0875	0.1013	0.1220	0.1729
24	0.0791	0.0875	0.1004	0.1221	0.1677
25	0.0784	0.0870	0.0939	0.1218	0.1710
26	0.0787	0.0883	0.1012	0.1204	0.1658
27	0.0794	0.0888	0.1001	0.1216	0.1719
28	0.0798	0.0892	0.1026	0.1239	0.1729
29	0.0791	0.0884	0.1006	0.1211	0.1747
30	0.0797	0.0881	0.1010	0.1241	0.1757
31	0.0794	0.0881	0.1001	0.1217	0.1671
32	0.0796	0.0873	0.1003	0.1221	0.1746
33	0.0797	0.0886	0.1012	0.1235	0.1687
34	0.0803	0.0891	0.1018	0.1227	0.1704
35	0.0797	0.0887	0.1010	0.1231	0.1751
36	0.0792	0.0888	0.1010	0.1236	0.1705
37	0.0799	0.0892	0.1021	0.1244	0.1698
38	0.0798	0.0891	0.1016	0.1221	0.1709
39	0.0801	0.0890	0.1017	0.1222	0.1694
40	0.0808	0.0895	0.1019	0.1234	0.1783
41	0.0797	0.0892	0.1029	0.1227	0.1732
42	0.0798	0.0885	0.1009	0.1239	0.1786
43	0.0790	0.0878	0.1007	0.1224	0.1746
44	0.0796	0.0894	0.1029	0.1253	0.1762
45	0.0804	0.0896	0.1008	0.1228	0.1731
46	0.0801	0.0888	0.1021	0.1219	0.1735
47	0.0805	0.0895	0.1017	0.1244	0.1778
48	0.0807	0.0898	0.1015	0.1226	0.1763
49	0.0810	0.0903	0.1022	0.1232	0.1771
50	0.0799	0.0890	0.1012	0.1254	0.1735

Table 4.5. Critical Values for the ANDERSON-DARLING Test with C = 0 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	1.0063	1.0353	1.0751	1.1356	1.2383
5	1.0934	1.1279	1.1721	1.2350	1.3968
6	1.1542	1.1892	1.2355	1.3114	1.4981
7	1.1965	1.2356	1.2856	1.3787	1.5673
8	1.2334	1.2728	1.3220	1.4069	1.6157
9	1.2635	1.3056	1.3567	1.4566	1.6869
10	1.2905	1.3328	1.3851	1.4858	1.6942
11	1.3065	1.3474	1.4041	1.5034	1.7557
12	1.3183	1.3608	1.4255	1.5349	1.7817
13	1.3313	1.3743	1.4322	1.5382	1.7897
14	1.3457	1.3888	1.4518	1.5684	1.8038
15	1.3570	1.4035	1.4656	1.5683	1.8315
16	1.3617	1.4031	1.4702	1.5836	1.8011
17	1.3770	1.4247	1.4874	1.5900	1.8182
18	1.3820	1.4273	1.4904	1.6026	1.8487
19	1.3845	1.4277	1.4885	1.6088	1.8504
20	1.3953	1.4381	1.5015	1.6060	1.8630
21	1.4035	1.4465	1.5147	1.6224	1.8736
22	1.4005	1.4428	1.5068	1.6265	1.8861
23	1.4063	1.4602	1.5243	1.6262	1.8897
24	1.4094	1.4574	1.5176	1.6250	1.8788
25	1.4128	1.4590	1.5307	1.6394	1.9106
26	1.4171	1.4626	1.5257	1.6280	1.8962
27	1.4217	1.4723	1.5389	1.6579	1.9217
28	1.4250	1.4734	1.5423	1.6512	1.9369
29	1.4298	1.4762	1.5381	1.6541	1.9198
30	1.4279	1.4767	1.5446	1.6563	1.9208
31	1.4330	1.4864	1.5534	1.6708	1.9408
32	1.4368	1.4882	1.5562	1.6703	1.9287
33	1.4366	1.4857	1.5511	1.6660	1.9226
34	1.4369	1.4840	1.5509	1.6651	1.9348
35	1.4457	1.4903	1.5593	1.6692	1.9444
36	1.4421	1.4900	1.5568	1.6712	1.9230
37	1.4403	1.4892	1.5519	1.6643	1.9237
38	1.4468	1.4932	1.5597	1.6763	1.9378
39	1.4509	1.4970	1.5632	1.6762	1.9363
40	1.4526	1.4997	1.5700	1.6848	1.9452
41	1.4485	1.4959	1.5647	1.6725	1.9472
42	1.4513	1.5001	1.5654	1.6757	1.9623
43	1.4509	1.5012	1.5711	1.6943	1.9544
44	1.4551	1.5011	1.5694	1.6937	1.9838
45	1.4574	1.5052	1.5707	1.6948	1.9789
46	1.4561	1.5013	1.5734	1.6814	1.9473
47	1.4549	1.5027	1.5766	1.6885	1.9553
48	1.4645	1.5110	1.5761	1.6894	1.9707
49	1.4553	1.5039	1.5747	1.6981	1.9642
50	1.4603	1.5116	1.5816	1.7020	1.9770

Table 4.6. Critical Values for the CRAMER-VON MISES Test with C = 0 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.0759	0.0813	0.0882	0.0997	0.1216
5	0.0769	0.0830	0.0905	0.1036	0.1313
6	0.0777	0.0845	0.0937	0.1094	0.1421
7	0.0779	0.0850	0.0951	0.1111	0.1458
8	0.0782	0.0852	0.0958	0.1123	0.1482
9	0.0782	0.0854	0.0951	0.1114	0.1491
10	0.0784	0.0855	0.0957	0.1141	0.1568
11	0.0795	0.0872	0.0987	0.1155	0.1595
12	0.0807	0.0881	0.0989	0.1172	0.1569
13	0.0798	0.0882	0.0992	0.1162	0.1562
14	0.0798	0.0881	0.0999	0.1180	0.1626
15	0.0805	0.0883	0.0993	0.1182	0.1644
16	0.0800	0.0881	0.0989	0.1205	0.1637
17	0.0798	0.0873	0.0993	0.1173	0.1665
18	0.0801	0.0883	0.0993	0.1204	0.1666
19	0.0800	0.0881	0.1003	0.1221	0.1711
20	0.0790	0.0878	0.1000	0.1200	0.1648
21	0.0804	0.0897	0.1008	0.1218	0.1641
22	0.0804	0.0886	0.1005	0.1194	0.1642
23	0.0796	0.0885	0.1012	0.1211	0.1694
24	0.0797	0.0885	0.1008	0.1213	0.1657
25	0.0796	0.0876	0.1002	0.1213	0.1664
26	0.0802	0.0890	0.1010	0.1200	0.1645
27	0.0804	0.0888	0.1004	0.1207	0.1686
28	0.0809	0.0901	0.1030	0.1242	0.1708
29	0.0796	0.0885	0.1012	0.1212	0.1729
30	0.0796	0.0887	0.1009	0.1230	0.1739
31	0.0802	0.0883	0.1000	0.1203	0.1657
32	0.0798	0.0881	0.1007	0.1197	0.1722
33	0.0801	0.0889	0.1016	0.1231	0.1649
34	0.0804	0.0893	0.1015	0.1224	0.1684
35	0.0805	0.0889	0.1016	0.1226	0.1725
36	0.0805	0.0890	0.1011	0.1235	0.1663
37	0.0810	0.0899	0.1022	0.1237	0.1673
38	0.0806	0.0894	0.1018	0.1215	0.1684
39	0.0807	0.0896	0.1022	0.1221	0.1689
40	0.0813	0.0898	0.1020	0.1232	0.1766
41	0.0799	0.0897	0.1028	0.1228	0.1695
42	0.0806	0.0894	0.1018	0.1234	0.1779
43	0.0796	0.0886	0.1009	0.1224	0.1736
44	0.0801	0.0896	0.1029	0.1257	0.1746
45	0.0807	0.0898	0.1014	0.1221	0.1690
46	0.0808	0.0892	0.1027	0.1221	0.1729
47	0.0810	0.0901	0.1021	0.1243	0.1768
48	0.0809	0.0901	0.1020	0.1227	0.1746
49	0.0816	0.0906	0.1027	0.1228	0.1767
50	0.0801	0.0891	0.1012	0.1250	0.1728

Table 4.7. Critical Values for the ANDERSON-DARLING Test with C = 0.375 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.5806	0.6154	0.6613	0.7342	0.8551
5	0.6185	0.6580	0.7079	0.7812	0.9685
6	0.6425	0.6825	0.7338	0.8206	1.0313
7	0.6556	0.6992	0.7551	0.8575	1.0704
8	0.6710	0.7133	0.7680	0.8599	1.0906
9	0.6831	0.7284	0.7838	0.8944	1.1465
10	0.6951	0.7410	0.7974	0.9051	1.1329
11	0.6991	0.7427	0.8041	0.9109	1.1852
12	0.6999	0.7454	0.8141	0.9318	1.1962
13	0.7031	0.7494	0.8106	0.9229	1.1911
14	0.7100	0.7556	0.8223	0.9445	1.1980
15	0.7152	0.7636	0.8279	0.9364	1.2196
16	0.7124	0.7574	0.8265	0.9479	1.1760
17	0.7235	0.7727	0.8387	0.9464	1.1860
18	0.7221	0.7696	0.8363	0.9544	1.2138
19	0.7208	0.7660	0.8304	0.9538	1.2065
20	0.7274	0.7717	0.8380	0.9472	1.2214
21	0.7322	0.7774	0.8487	0.9597	1.2225
22	0.7246	0.7696	0.8370	0.9588	1.2296
23	0.7284	0.7854	0.8500	0.9546	1.2320
24	0.7284	0.7775	0.8411	0.9532	1.2151
25	0.7293	0.7776	0.8520	0.9638	1.2471
26	0.7311	0.7783	0.8426	0.9484	1.2268
27	0.7341	0.7861	0.8527	0.9775	1.2497
28	0.7344	0.7857	0.8553	0.9684	1.2639
29	0.7380	0.7856	0.8500	0.9689	1.2421
30	0.7342	0.7854	0.8541	0.9683	1.2418
31	0.7374	0.7920	0.8601	0.9829	1.2588
32	0.7390	0.7923	0.8617	0.9797	1.2458
33	0.7383	0.7879	0.8565	0.9724	1.2359
34	0.7356	0.7857	0.8535	0.9695	1.2465
35	0.7436	0.7894	0.8600	0.9758	1.2553
36	0.7379	0.7881	0.8565	0.9760	1.2322
37	0.7358	0.7858	0.8503	0.9651	1.2362
38	0.7421	0.7888	0.8556	0.9750	1.2472
39	0.7448	0.7914	0.8592	0.9764	1.2437
40	0.7449	0.7927	0.8659	0.9843	1.2517
41	0.7399	0.7886	0.8591	0.9687	1.2508
42	0.7418	0.7922	0.8579	0.9716	1.2625
43	0.7405	0.7919	0.8635	0.9889	1.2537
44	0.7438	0.7914	0.8600	0.9871	1.2815
45	0.7456	0.7939	0.8612	0.9886	1.2772
46	0.7432	0.7902	0.8624	0.9730	1.2417
47	0.7408	0.7903	0.8643	0.9799	1.2519
48	0.7506	0.7979	0.8640	0.9792	1.2659
49	0.7393	0.7891	0.8618	0.9872	1.2616
50	0.7445	0.7968	0.8673	0.9908	1.2766

Table 4.8. Critical Values for the CRAMER-VON MISES Test with C = 0.375 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.0692	0.0757	0.0847	0.1010	0.1283
5	0.0721	0.0794	0.0887	0.1042	0.1442
6	0.0740	0.0818	0.0927	0.1114	0.1532
7	0.0752	0.0829	0.0942	0.1125	0.1563
8	0.0757	0.0835	0.0948	0.1145	0.1578
9	0.0752	0.0833	0.0953	0.1121	0.1578
10	0.0760	0.0839	0.0955	0.1160	0.1658
11	0.0777	0.0863	0.0981	0.1171	0.1662
12	0.0784	0.0873	0.0981	0.1196	0.1647
13	0.0783	0.0869	0.0987	0.1182	0.1608
14	0.0778	0.0875	0.0994	0.1191	0.1701
15	0.0787	0.0872	0.0995	0.1191	0.1699
16	0.0786	0.0866	0.0994	0.1223	0.1694
17	0.0786	0.0868	0.0982	0.1186	0.1695
18	0.0789	0.0879	0.1000	0.1211	0.1729
19	0.0789	0.0878	0.0996	0.1231	0.1788
20	0.0782	0.0867	0.1003	0.1206	0.1688
21	0.0795	0.0887	0.1011	0.1227	0.1683
22	0.0799	0.0881	0.1000	0.1209	0.1663
23	0.0788	0.0878	0.1015	0.1224	0.1738
24	0.0794	0.0878	0.1008	0.1227	0.1682
25	0.0786	0.0873	0.1001	0.1223	0.1712
26	0.0789	0.0883	0.1016	0.1211	0.1669
27	0.0794	0.0889	0.1009	0.1219	0.1727
28	0.0799	0.0895	0.1030	0.1244	0.1739
29	0.0795	0.0883	0.1011	0.1213	0.1748
30	0.0798	0.0883	0.1011	0.1244	0.1755
31	0.0795	0.0884	0.1002	0.1217	0.1676
32	0.0796	0.0876	0.1009	0.1226	0.1758
33	0.0798	0.0887	0.1013	0.1239	0.1696
34	0.0803	0.0892	0.1022	0.1228	0.1717
35	0.0797	0.0889	0.1014	0.1233	0.1758
36	0.0794	0.0890	0.1011	0.1239	0.1715
37	0.0800	0.0894	0.1025	0.1245	0.1708
38	0.0798	0.0892	0.1018	0.1226	0.1716
39	0.0801	0.0892	0.1017	0.1227	0.1700
40	0.0807	0.0897	0.1022	0.1238	0.1784
41	0.0801	0.0895	0.1029	0.1229	0.1738
42	0.0799	0.0887	0.1010	0.1238	0.1784
43	0.0791	0.0880	0.1009	0.1225	0.1757
44	0.0797	0.0895	0.1030	0.1253	0.1762
45	0.0806	0.0896	0.1012	0.1231	0.1735
46	0.0800	0.0889	0.1026	0.1225	0.1743
47	0.0805	0.0896	0.1020	0.1251	0.1779
48	0.0808	0.0897	0.1017	0.1226	0.1769
49	0.0809	0.0903	0.1024	0.1233	0.1777
50	0.0799	0.0891	0.1014	0.1255	0.1738

Table 4.9. Critical Values for the ANDERSON-DARLING Test with C = 0.3 (Standard) Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.6775	0.7108	0.7554	0.8254	0.9422
5	0.7240	0.7622	0.8113	0.8818	1.0637
6	0.7545	0.7934	0.8435	0.9278	1.1336
7	0.7722	0.8153	0.8701	0.9698	1.1776
8	0.7918	0.8331	0.8864	0.9772	1.2028
9	0.8067	0.8512	0.9059	1.0142	1.2615
10	0.8210	0.8662	0.9217	1.0275	1.2519
11	0.8272	0.8701	0.9302	1.0356	1.3056
12	0.8291	0.8747	0.9426	1.0582	1.3196
13	0.8344	0.8798	0.9405	1.0509	1.3163
14	0.8423	0.8876	0.9528	1.0746	1.3238
15	0.8484	0.8963	0.9604	[illegible]	1.3471
16	0.8472	0.8909	0.9604	1.0800	1.3057
17	0.8583	0.9074	0.9735	1.0794	1.3182
18	0.8587	0.9058	0.9718	1.0876	1.3449
19	0.8576	0.9023	0.9658	1.0888	1.3385
20	0.8646	0.9090	0.9742	1.0827	1.3535
21	0.8704	0.9153	0.9854	1.0965	1.3567
22	0.8634	0.9077	0.9747	1.0959	1.3644
23	0.8678	0.9236	0.9877	1.0918	1.3667
24	0.8681	0.9168	0.9796	1.0911	1.3517
25	0.8692	0.9167	0.9910	1.1020	1.3840
26	0.8713	0.9185	0.9824	1.0873	1.3635
27	0.8740	0.9265	0.9930	1.1161	1.3853
28	0.8752	0.9263	0.9952	1.1076	1.4022
29	0.8790	0.9267	0.9903	1.1085	1.3799
30	0.8756	0.9260	0.9954	1.1081	1.3806
31	0.8793	0.9334	1.0012	1.1232	1.3977
32	0.8812	0.9344	1.0032	1.1202	1.3849
33	0.8800	0.9300	0.9984	1.1138	1.3759
34	0.8782	0.9277	0.9950	1.1110	1.3873
35	0.8861	0.9323	1.0026	1.1166	1.3952
36	0.8812	0.9305	0.9990	1.1181	1.3720
37	0.8792	0.9288	0.9931	1.1071	1.3759
38	0.8852	0.9317	0.9987	1.1175	1.3881
39	0.8882	0.9346	1.0024	1.1187	1.3828
40	0.8886	0.9361	1.0089	1.1264	1.3928
41	0.8838	0.9321	1.0024	1.1116	1.3919
42	0.8859	0.9358	1.0016	1.1151	1.4045
43	0.8843	0.9358	1.0071	1.1316	1.3958
44	0.8879	0.9354	1.0039	1.1307	1.4241
45	0.8896	0.9378	1.0048	1.1319	1.4196
46	0.8873	0.9341	1.0060	1.1164	1.3837
47	0.8855	0.9345	1.0086	1.1230	1.3943
48	0.8951	0.9425	1.0084	1.1227	1.4091
49	0.8840	0.9341	1.0062	1.1308	1.4037
50	0.8892	0.9414	1.0117	1.1347	1.4181

Table 4.10. Critical Values for the CRAMER-VON MISES Test with C = 0.3 (Standard) Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.0696	0.0759	0.0843	0.0993	0.1257
5	0.0721	0.0792	0.0880	0.1031	0.1403
6	0.0741	0.0815	0.0913	0.1099	0.1500
7	0.0749	0.0825	0.0936	0.1111	0.1526
8	0.0754	0.0831	0.0940	0.1133	0.1544
9	0.0752	0.0831	0.0946	0.1112	0.1552
10	0.0758	0.0836	0.0945	0.1146	0.1628
11	0.0774	0.0856	0.0972	0.1161	0.1630
12	0.0784	0.0869	0.0977	0.1186	0.1623
13	0.0781	0.0866	0.0983	0.1169	0.1592
14	0.0778	0.0871	0.0990	0.1182	0.1680
15	0.0783	0.0871	0.0989	0.1180	0.1684
16	0.0785	0.0865	0.0988	0.1215	0.1678
17	0.0785	0.0865	0.0977	0.1179	0.1683
18	0.0788	0.0875	0.0995	0.1211	0.1714
19	0.0788	0.0874	0.0995	0.1221	0.1770
20	0.0781	0.0866	0.0997	0.1200	0.1676
21	0.0792	0.0884	0.1006	0.1222	0.1663
22	0.0795	0.0878	0.0998	0.1203	0.1646
23	0.0789	0.0875	0.1014	0.1217	0.1721
24	0.0791	0.0876	0.1003	0.1219	0.1670
25	0.0783	0.0870	0.0997	0.1214	0.1706
26	0.0786	0.0884	0.1011	0.1203	0.1657
27	0.0794	0.0888	0.1001	0.1214	0.1715
28	0.0797	0.0891	0.1024	0.1238	0.1728
29	0.0791	0.0884	0.1006	0.1209	0.1749
30	0.0796	0.0881	0.1009	0.1241	0.1756
31	0.0793	0.0879	0.1000	0.1217	0.1669
32	0.0796	0.0872	0.1001	0.1219	0.1743
33	0.0796	0.0884	0.1012	0.1234	0.1685
34	0.0802	0.0891	0.1018	0.1225	0.1701
35	0.0796	0.0887	0.1010	0.1230	0.1745
36	0.0792	0.0888	0.1010	0.1237	0.1701
37	0.0799	0.0892	0.1021	0.1242	0.1696
38	0.0799	0.0891	0.1015	0.1220	0.1704
39	0.0801	0.0892	0.1017	0.1222	0.1692
40	0.0807	0.0895	0.1018	0.1231	0.1779
41	0.0797	0.0892	0.1029	0.1227	0.1728
42	0.0799	0.0886	0.1009	0.1239	0.1787
43	0.0788	0.0879	0.1006	0.1224	0.1741
44	0.0797	0.0894	0.1029	0.1252	0.1764
45	0.0804	0.0896	0.1007	0.1228	0.1727
46	0.0801	0.0888	0.1020	0.1219	0.1736
47	0.0805	0.0894	0.1018	0.1243	0.1779
48	0.0807	0.0898	0.1014	0.1226	0.1763
49	0.0810	0.0903	0.1022	0.1231	0.1772
50	0.0799	0.0889	0.1010	0.1253	0.1734

Table 4.11. Critical Values for the ANDERSON-DARLING Test with C = 1 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	-.6018	-.5521	-.4856	-.3810	-.2089
5	-.5695	-.5180	-.4502	-.3544	-.1020
6	-.5521	-.5036	-.4379	-.3264	-.0546
7	-.5460	-.4933	-.4236	-.2998	-.0323
8	-.5343	-.4842	-.4175	-.3105	-.0289
9	-.5256	-.4738	-.4065	-.2786	0.0185
10	-.5177	-.4639	-.3973	-.2757	-.0105
11	-.5153	-.4668	-.3986	-.2749	0.0334
12	-.5186	-.4670	-.3903	-.2552	0.0388
13	-.5179	-.4652	-.3986	-.2709	0.0279
14	-.5112	-.4613	-.3885	-.2520	0.0363
15	-.5091	-.4535	-.3823	-.2632	0.0462
16	-.5134	-.4650	-.3887	-.2544	0.0011
17	-.5031	-.4499	-.3769	-.2594	-.0049
18	-.5061	-.4548	-.3820	-.2525	0.0348
19	-.5079	-.4600	-.3910	-.2553	0.0184
20	-.5027	-.4525	-.3834	-.2649	0.0299
21	-.4978	-.4495	-.3724	-.2518	0.0292
22	-.5065	-.4594	-.3887	-.2563	0.0326
23	-.5036	-.4432	-.3739	-.2614	0.0323
24	-.5050	-.4534	-.3839	-.2645	0.0085
25	-.5034	-.4532	-.3757	-.2552	0.0387
26	-.5027	-.4517	-.3837	-.2728	0.0208
27	-.4991	-.4449	-.3743	-.2442	0.0426
28	-.4997	-.4463	-.3700	-.2520	0.0550
29	-.4978	-.4487	-.3786	-.2547	0.0352
30	-.5024	-.4475	-.3762	-.2550	0.0287
31	-.4986	-.4424	-.3708	-.2415	0.0478
32	-.4977	-.4419	-.3692	-.2455	0.0350
33	-.4992	-.4472	-.3767	-.2540	0.0194
34	-.5022	-.4498	-.3785	-.2589	0.0310
35	-.4946	-.4469	-.3722	-.2519	0.0429
36	-.5001	-.4472	-.3767	-.2515	0.0173
37	-.5025	-.4499	-.3832	-.2632	0.0238
38	-.4974	-.4482	-.3786	-.2554	0.0298
39	-.4948	-.4464	-.3761	-.2534	0.0266
40	-.4947	-.4451	-.3701	-.2465	0.0300
41	-.4997	-.4497	-.3756	-.2630	0.0353
42	-.4994	-.4471	-.3780	-.2613	0.0453
43	-.4991	-.4473	-.3719	-.2434	0.0370
44	-.4966	-.4476	-.3750	-.2442	0.0640
45	-.4938	-.4453	-.3745	-.2458	0.0533
46	-.4964	-.4497	-.3746	-.2562	0.0196
47	-.5005	-.4482	-.3719	-.2543	0.0277
48	-.4904	-.4413	-.3735	-.2543	0.0361
49	-.5027	-.4506	-.3758	-.2447	0.0394
50	-.4968	-.4427	-.3721	-.2417	0.0593

Table 4.12. Critical Values for the CRAMER-VON MISES Test with C = 1 Median Rank Plotting Position; alpha = .20 ... .01

N	.20	.15	.10	.05	.01
4	0.1185	0.1296	0.1461	0.1730	0.2123
5	0.1118	0.1234	0.1383	0.1645	0.2267
6	0.1082	0.1199	0.1357	0.1630	0.2304
7	0.1044	0.1152	0.1312	0.1587	0.2239
8	0.1017	0.1127	0.1272	0.1556	0.2175
9	0.0977	0.1087	0.1238	0.1489	0.2085
10	0.0968	0.1071	0.1222	0.1483	0.2128
11	0.0964	0.1071	0.1225	0.1476	0.2106
12	0.0958	0.1062	0.1213	0.1478	0.2036
13	0.0948	0.1049	0.1202	0.1458	0.2029
14	0.0929	0.1035	0.1175	0.1440	0.2027
15	0.0926	0.1030	0.1168	0.1425	0.2040
16	0.0909	0.1009	0.1168	0.1427	0.2036
17	0.0910	0.1011	0.1139	0.1383	0.2009
18	0.0903	0.1012	0.1161	0.1413	0.2010
19	0.0903	0.1000	0.1143	0.1404	0.2051
20	0.0882	0.0984	0.1134	0.1367	0.1933
21	0.0896	0.0995	0.1143	0.1392	0.1951
22	0.0891	0.0988	0.1125	0.1381	0.1936
23	0.0886	0.0985	0.1131	0.1376	0.1930
24	0.0885	0.0986	0.1128	0.1360	0.1903
25	0.0872	0.0970	0.1106	0.1355	0.1934
26	0.0869	0.0971	0.1124	0.1359	0.1861
27	0.0869	0.0970	0.1108	0.1346	0.1924
28	0.0878	0.0989	0.1130	0.1362	0.1915
29	0.0868	0.0973	0.1102	0.1329	0.1944
30	0.0866	0.0968	0.1110	0.1358	0.1946
31	0.0867	0.0958	0.1093	0.1330	0.1835
32	0.0864	0.0954	0.1084	0.1334	0.1899
33	0.0867	0.0958	0.1099	0.1344	0.1840
34	0.0866	0.0971	0.1102	0.1329	0.1875
35	0.0863	0.0960	0.1098	0.1336	0.1913
36	0.0857	0.0955	0.1086	0.1327	0.1887
37	0.0854	0.0957	0.1089	0.1336	0.1827
38	0.0855	0.0955	0.1094	0.1322	0.1844
39	0.0858	0.0951	0.1078	0.1317	0.1824
40	0.0858	0.0961	0.1092	0.1328	0.1919
41	0.0855	0.0954	0.1088	0.1318	0.1866
42	0.0850	0.0943	0.1075	0.1301	0.1881
43	0.0836	0.0937	0.1071	0.1305	0.1857
44	0.0848	0.0949	0.1095	0.1336	0.1881
45	0.0857	0.0951	0.1075	0.1307	0.1832
46	0.0847	0.0946	0.1081	0.1298	0.1872
47	0.0848	0.0941	0.1084	0.1338	0.1919
48	0.0852	0.0948	0.1069	0.1309	0.1883
49	0.0853	0.0944	0.1079	0.1311	0.1887
50	0.0841	0.0937	0.1076	0.1336	0.1853

4.2 Power Study Tables

To use these tables, select the α level from across the top row, and go down the table to the cluster of sample sizes of interest to compare the six different powers.

Table 4.13. ANDERSON-DARLING GOFT's Power Against UNIFORM(0,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.285	0.220	0.158	0.079	0.016	Unmodified
10	0.286	0.221	0.161	0.081	0.016	C = 0.3175 Median Rank
10	0.290	0.224	0.164	0.082	0.017	C = 0 Median Rank
10	0.285	0.220	0.161	0.080	0.016	C = 0.375 Median Rank
10	0.286	0.221	0.161	0.081	0.016	C = 0.3 (Standard) Median Rank
10	0.281	0.212	0.151	0.076	0.015	C = 1 Median Rank
20	0.464	0.392	0.300	0.176	0.040	Unmodified
20	0.468	0.396	0.303	0.180	0.040	C = 0.3175 Median Rank
20	0.474	0.401	0.308	0.183	0.043	C = 0 Median Rank
20	0.466	0.395	0.302	0.178	0.040	C = 0.375 Median Rank
20	0.469	0.396	0.303	0.180	0.041	C = 0.3 (Standard) Median Rank
20	0.454	0.381	0.289	0.170	0.037	C = 1 Median Rank
30	0.632	0.553	0.449	0.299	0.085	Unmodified
30	0.637	0.556	0.451	0.302	0.087	C = 0.3175 Median Rank
30	0.644	0.562	0.458	0.307	0.091	C = 0 Median Rank
30	0.636	0.554	0.451	0.301	0.087	C = 0.375 Median Rank
30	0.637	0.556	0.452	0.302	0.087	C = 0.3 (Standard) Median Rank
30	0.625	0.542	0.439	0.289	0.082	C = 1 Median Rank
40	0.765	0.698	0.594	0.431	0.169	Unmodified
40	0.767	0.700	0.596	0.435	0.172	C = 0.3175 Median Rank
40	0.772	0.707	0.603	0.445	0.179	C = 0 Median Rank
40	0.767	0.700	0.595	0.433	0.170	C = 0.375 Median Rank
40	0.768	0.701	0.596	0.436	0.172	C = 0.3 (Standard) Median Rank
40	0.759	0.690	0.586	0.424	0.166	C = 1 Median Rank
50	0.862	0.808	0.723	0.564	0.260	Unmodified
50	0.864	0.810	0.726	0.567	0.263	C = 0.3175 Median Rank
50	0.867	0.817	0.729	0.574	0.271	C = 0 Median Rank
50	0.863	0.810	0.726	0.566	0.262	C = 0.375 Median Rank
50	0.864	0.810	0.726	0.568	0.264	C = 0.3 (Standard) Median Rank
50	0.857	0.800	0.719	0.552	0.246	C = 1 Median Rank

Table 4.14. ANDERSON-DARLING GOFT's Power Against BETA(3,2) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.219	0.162	0.108	0.053	0.012	Unmodified
10	0.219	0.163	0.108	0.053	0.012	C = 0.3175 Median Rank
10	0.220	0.163	0.110	0.052	0.013	C = 0 Median Rank
10	0.219	0.163	0.108	0.053	0.012	C = 0.375 Median Rank
10	0.219	0.163	0.109	0.053	0.012	C = 0.3 (Standard) Median Rank
10	0.216	0.159	0.106	0.053	0.012	C = 1 Median Rank
20	0.256	0.192	0.133	0.072	0.014	Unmodified
20	0.258	0.194	0.134	0.072	0.014	C = 0.3175 Median Rank
20	0.259	0.197	0.136	0.073	0.015	C = 0 Median Rank
20	0.257	0.194	0.133	0.072	0.014	C = 0.375 Median Rank
20	0.258	0.194	0.134	0.072	0.014	C = 0.3 (Standard) Median Rank
20	0.253	0.186	0.130	0.070	0.014	C = 1 Median Rank
30	0.322	0.249	0.180	0.101	0.022	Unmodified
30	0.327	0.251	0.182	0.102	0.023	C = 0.3175 Median Rank
30	0.330	0.253	0.185	0.103	0.024	C = 0 Median Rank
30	0.326	0.250	0.182	0.102	0.023	C = 0.375 Median Rank
30	0.327	0.251	0.182	0.102	0.023	C = 0.3 (Standard) Median Rank
30	0.317	0.243	0.177	0.098	0.022	C = 1 Median Rank
40	0.385	0.318	0.226	0.132	0.031	Unmodified
40	0.386	0.320	0.228	0.132	0.032	C = 0.3175 Median Rank
40	0.388	0.323	0.232	0.135	0.034	C = 0 Median Rank
40	0.386	0.319	0.227	0.132	0.031	C = 0.375 Median Rank
40	0.386	0.320	0.228	0.133	0.032	C = 0.3 (Standard) Median Rank
40	0.381	0.312	0.223	0.128	0.030	C = 1 Median Rank
50	0.453	0.373	0.284	0.163	0.040	Unmodified
50	0.456	0.376	0.287	0.163	0.040	C = 0.3175 Median Rank
50	0.459	0.379	0.289	0.166	0.042	C = 0 Median Rank
50	0.455	0.376	0.286	0.162	0.040	C = 0.375 Median Rank
50	0.456	0.376	0.286	0.163	0.040	C = 0.3 (Standard) Median Rank
50	0.447	0.367	0.282	0.156	0.036	C = 1 Median Rank

Table 4.15. ANDERSON-DARLING GOFT's Power Against EXPONENTIAL(1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.668	0.605	0.525	0.399	0.215	Unmodified
10	0.668	0.604	0.527	0.399	0.214	C = 0.3175 Median Rank
10	0.667	0.604	0.526	0.397	0.215	C = 0 Median Rank
10	0.668	0.604	0.526	0.399	0.214	C = 0.375 Median Rank
10	0.668	0.604	0.527	0.399	0.214	C = 0.3 (Standard) Median Rank
10	0.666	0.603	0.525	0.399	0.215	C = 1 Median Rank
20	0.925	0.902	0.862	0.783	0.560	Unmodified
20	0.925	0.902	0.863	0.783	0.561	C = 0.3175 Median Rank
20	0.926	0.903	0.864	0.783	0.565	C = 0 Median Rank
20	0.925	0.903	0.862	0.784	0.560	C = 0.375 Median Rank
20	0.925	0.902	0.863	0.783	0.561	C = 0.3 (Standard) Median Rank
20	0.926	0.901	0.863	0.782	0.563	C = 1 Median Rank
30	0.984	0.978	0.965	0.935	0.819	Unmodified
30	0.985	0.978	0.965	0.935	0.819	C = 0.3175 Median Rank
30	0.985	0.978	0.965	0.934	0.820	C = 0 Median Rank
30	0.985	0.978	0.965	0.935	0.819	C = 0.375 Median Rank
30	0.985	0.978	0.965	0.935	0.819	C = 0.3 (Standard) Median Rank
30	0.984	0.977	0.965	0.936	0.819	C = 1 Median Rank
40	0.998	0.996	0.992	0.983	0.935	Unmodified
40	0.998	0.996	0.992	0.983	0.935	C = 0.3175 Median Rank
40	0.998	0.996	0.992	0.983	0.936	C = 0 Median Rank
40	0.998	0.996	0.992	0.983	0.935	C = 0.375 Median Rank
40	0.998	0.996	0.992	0.983	0.935	C = 0.3 (Standard) Median Rank
40	0.998	0.995	0.992	0.983	0.935	C = 1 Median Rank
50	1.000	1.000	0.999	0.995	0.978	Unmodified
50	1.000	1.000	0.999	0.995	0.979	C = 0.3175 Median Rank
50	1.000	1.000	0.999	0.995	0.980	C = 0 Median Rank
50	1.000	1.000	0.999	0.995	0.979	C = 0.375 Median Rank
50	1.000	1.000	0.999	0.995	0.979	C = 0.3 (Standard) Median Rank
50	1.000	1.000	0.999	0.995	0.978	C = 1 Median Rank

Table 4.16. ANDERSON-DARLING GOFT's Power Against DBL EXPONENTIAL(2,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.721	0.667	0.596	0.485	0.291	Unmodified
10	0.722	0.667	0.598	0.485	0.290	C = 0.3175 Median Rank
10	0.720	0.667	0.598	0.482	0.290	C = 0 Median Rank
10	0.722	0.667	0.598	0.485	0.291	C = 0.375 Median Rank
10	0.721	0.667	0.598	0.485	0.290	C = 0.3 (Standard) Median Rank
10	0.721	0.666	0.597	0.486	0.290	C = 1 Median Rank
20	0.958	0.942	0.911	0.857	0.687	Unmodified
20	0.958	0.942	0.912	0.858	0.689	C = 0.3175 Median Rank
20	0.958	0.942	0.911	0.859	0.692	C = 0 Median Rank
20	0.958	0.942	0.911	0.858	0.688	C = 0.375 Median Rank
20	0.958	0.942	0.912	0.858	0.689	C = 0.3 (Standard) Median Rank
20	0.957	0.942	0.913	0.857	0.688	C = 1 Median Rank
30	0.994	0.990	0.984	0.970	0.908	Unmodified
30	0.995	0.990	0.984	0.970	0.908	C = 0.3175 Median Rank
30	0.995	0.991	0.984	0.970	0.908	C = 0 Median Rank
30	0.995	0.990	0.984	0.970	0.908	C = 0.375 Median Rank
30	0.995	0.990	0.984	0.970	0.908	C = 0.3 (Standard) Median Rank
30	0.994	0.990	0.984	0.969	0.908	C = 1 Median Rank
40	1.000	0.999	0.997	0.993	0.974	Unmodified
40	1.000	0.999	0.997	0.993	0.974	C = 0.3175 Median Rank
40	1.000	0.999	0.997	0.994	0.974	C = 0 Median Rank
40	1.000	0.999	0.997	0.993	0.974	C = 0.375 Median Rank
40	1.000	0.999	0.997	0.993	0.974	C = 0.3 (Standard) Median Rank
40	1.000	0.999	0.997	0.993	0.974	C = 1 Median Rank
50	1.000	1.000	1.000	0.999	0.994	Unmodified
50	1.000	1.000	1.000	0.999	0.994	C = 0.3175 Median Rank
50	1.000	1.000	1.000	0.999	0.994	C = 0 Median Rank
50	1.000	1.000	1.000	0.999	0.994	C = 0.375 Median Rank
50	1.000	1.000	1.000	0.999	0.994	C = 0.3 (Standard) Median Rank
50	1.000	1.000	1.000	0.999	0.993	C = 1 Median Rank

Table 4.17. ANDERSON-DARLING GOFT's Power Against WEIBULL(2,6) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.243	0.190	0.139	0.075	0.018	Unmodified
10	0.243	0.190	0.140	0.075	0.018	C = 0.3175 Median Rank
10	0.243	0.189	0.140	0.074	0.018	C = 0 Median Rank
10	0.243	0.189	0.140	0.075	0.018	C = 0.375 Median Rank
10	0.243	0.190	0.140	0.075	0.018	C = 0.3 (Standard) Median Rank
10	0.245	0.190	0.139	0.076	0.018	C = 1 Median Rank
20	0.357	0.288	0.215	0.131	0.037	Unmodified
20	0.357	0.288	0.214	0.131	0.037	C = 0.3175 Median Rank
20	0.358	0.288	0.214	0.130	0.037	C = 0 Median Rank
20	0.357	0.289	0.214	0.131	0.037	C = 0.375 Median Rank
20	0.357	0.288	0.214	0.131	0.037	C = 0.3 (Standard) Median Rank
20	0.356	0.284	0.213	0.130	0.037	C = 1 Median Rank
30	0.454	0.373	0.284	0.182	0.058	Unmodified
30	0.455	0.373	0.283	0.182	0.059	C = 0.3175 Median Rank
30	0.457	0.375	0.284	0.182	0.060	C = 0 Median Rank
30	0.455	0.372	0.283	0.182	0.059	C = 0.375 Median Rank
30	0.455	0.373	0.283	0.182	0.059	C = 0.3 (Standard) Median Rank
30	0.455	0.368	0.283	0.180	0.060	C = 1 Median Rank
40	0.526	0.454	0.362	0.241	0.091	Unmodified
40	0.526	0.454	0.363	0.242	0.091	C = 0.3175 Median Rank
40	0.527	0.453	0.364	0.244	0.092	C = 0 Median Rank
40	0.526	0.454	0.362	0.241	0.091	C = 0.375 Median Rank
40	0.526	0.454	0.363	0.242	0.091	C = 0.3 (Standard) Median Rank
40	0.523	0.453	0.363	0.240	0.091	C = 1 Median Rank
50	0.611	0.531	0.439	0.305	0.115	Unmodified
50	0.612	0.532	0.440	0.304	0.116	C = 0.3175 Median Rank
50	0.612	0.532	0.440	0.305	0.118	C = 0 Median Rank
50	0.611	0.531	0.441	0.304	0.116	C = 0.375 Median Rank
50	0.612	0.532	0.440	0.304	0.116	C = 0.3 (Standard) Median Rank
50	0.609	0.531	0.439	0.301	0.111	C = 1 Median Rank

Table 4.18. ANDERSON-DARLING GOFT's Power Against LOG-NORMAL(0,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.779	0.733	0.677	0.583	0.406	Unmodified
10	0.779	0.732	0.677	0.583	0.406	C = 0.3175 Median Rank
10	0.779	0.730	0.678	0.582	0.406	C = 0 Median Rank
10	0.779	0.732	0.677	0.583	0.406	C = 0.375 Median Rank
10	0.779	0.732	0.677	0.582	0.406	C = 0.3 (Standard) Median Rank
10	0.779	0.732	0.676	0.583	0.408	C = 1 Median Rank
20	0.974	0.963	0.944	0.905	0.806	Unmodified
20	0.975	0.962	0.944	0.905	0.806	C = 0.3175 Median Rank
20	0.975	0.962	0.944	0.905	0.808	C = 0 Median Rank
20	0.974	0.963	0.944	0.905	0.806	C = 0.375 Median Rank
20	0.975	0.962	0.944	0.905	0.806	C = 0.3 (Standard) Median Rank
20	0.974	0.962	0.944	0.905	0.807	C = 1 Median Rank
30	0.998	0.995	0.992	0.983	0.950	Unmodified
30	0.998	0.995	0.992	0.983	0.950	C = 0.3175 Median Rank
30	0.998	0.995	0.992	0.983	0.950	C = 0 Median Rank
30	0.998	0.995	0.992	0.983	0.950	C = 0.375 Median Rank
30	0.998	0.995	0.992	0.983	0.950	C = 0.3 (Standard) Median Rank
30	0.998	0.995	0.992	0.983	0.951	C = 1 Median Rank
40	1.000	1.000	0.999	0.998	0.990	Unmodified
40	1.000	1.000	0.999	0.998	0.990	C = 0.3175 Median Rank
40	1.000	1.000	0.999	0.998	0.991	C = 0 Median Rank
40	1.000	1.000	0.999	0.998	0.990	C = 0.375 Median Rank
40	1.000	1.000	0.999	0.998	0.990	C = 0.3 (Standard) Median Rank
40	1.000	1.000	0.999	0.998	0.990	C = 1 Median Rank
50	1.000	1.000	1.000	1.000	0.999	Unmodified
50	1.000	1.000	1.000	1.000	0.999	C = 0.3175 Median Rank
50	1.000	1.000	1.000	1.000	0.999	C = 0 Median Rank
50	1.000	1.000	1.000	1.000	0.999	C = 0.375 Median Rank
50	1.000	1.000	1.000	1.000	0.999	C = 0.3 (Standard) Median Rank
50	1.000	1.000	1.000	1.000	0.999	C = 1 Median Rank

Table 4.19. CRAMER-VON MISES GOFT's Power Against UNIFORM(0,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.285	0.226	0.156	0.079	0.012	Unmodified
10	0.306	0.250	0.175	0.092	0.015	C = 0.3175 Median Rank
10	0.349	0.285	0.208	0.115	0.023	C = 0 Median Rank
10	0.303	0.242	0.171	0.088	0.014	C = 0.375 Median Rank
10	0.311	0.252	0.179	0.093	0.016	C = 0.3 (Standard) Median Rank
10	0.220	0.164	0.107	0.052	0.008	C = 1 Median Rank
20	0.426	0.347	0.257	0.146	0.034	Unmodified
20	0.459	0.382	0.285	0.164	0.039	C = 0.3175 Median Rank
20	0.521	0.438	0.334	0.204	0.054	C = 0 Median Rank
20	0.449	0.373	0.273	0.158	0.037	C = 0.375 Median Rank
20	0.463	0.384	0.288	0.168	0.040	C = 0.3 (Standard) Median Rank
20	0.329	0.257	[illegible]74	0.090	0.018	C = 1 Median Rank
30	0.552	0.471	0.372	0.214	0.058	Unmodified
30	0.585	0.512	0.406	0.243	0.066	C = 0.3175 Median Rank
30	0.647	0.566	0.460	0.295	0.086	C = 0 Median Rank
30	0.574	0.499	0.395	0.234	0.065	C = 0.375 Median Rank
30	0.589	0.517	0.408	0.245	0.066	C = 0.3 (Standard) Median Rank
30	0.447	0.366	0.261	0.147	0.030	C = 1 Median Rank
40	0.667	0.584	0.486	0.330	0.106	Unmodified
40	0.696	0.621	0.523	0.365	0.118	C = 0.3175 Median Rank
40	0.742	0.678	0.575	0.413	0.143	C = 0 Median Rank
40	0.688	0.611	0.511	0.356	0.115	C = 0.375 Median Rank
40	0.699	0.623	0.526	0.368	0.120	C = 0.3 (Standard) Median Rank
40	0.569	0.486	0.382	0.243	0.063	C = 1 Median Rank
50	0.768	0.701	0.603	0.426	0.176	Unmodified
50	0.789	0.728	0.636	0.460	0.196	C = 0.3175 Median Rank
50	0.826	0.772	0.686	0.511	0.237	C = 0 Median Rank
50	0.783	0.720	0.626	0.453	0.191	C = 0.375 Median Rank
50	0.790	0.731	0.638	0.464	0.198	C = 0.3 (Standard) Median Rank
50	0.686	0.609	0.503	0.329	0.120	C = 1 Median Rank

Table 4.20. CRAMER-VON MISES GOFT's Power Against BETA(3,2) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.226	0.176	0.115	0.056	0.010	Unmodified
10	0.241	0.185	0.123	0.061	0.010	C = 0.3175 Median Rank
10	0.258	0.199	0.138	0.068	0.011	C = 0 Median Rank
10	0.237	0.182	0.120	0.059	0.010	C = 0.375 Median Rank
10	0.243	0.186	0.126	0.061	0.010	C = 0.3 (Standard) Median Rank
10	0.202	0.147	0.097	0.051	0.009	C = 1 Median Rank
20	0.260	0.199	0.130	0.072	0.015	Unmodified
20	0.271	0.213	0.139	0.075	0.017	C = 0.3175 Median Rank
20	0.301	0.233	0.160	0.083	0.021	C = 0 Median Rank
20	0.268	0.208	0.136	0.074	0.016	C = 0.375 Median Rank
20	0.273	0.213	0.139	0.076	0.017	C = 0.3 (Standard) Median Rank
20	0.217	0.159	0.101	0.057	0.011	C = 1 Median Rank
30	0.300	0.239	0.166	0.089	0.020	Unmodified
30	0.319	0.252	0.178	0.097	0.021	C = 0.3175 Median Rank
30	0.356	0.281	0.202	0.112	0.024	C = 0 Median Rank
30	0.312	0.248	0.175	0.094	0.022	C = 0.375 Median Rank
30	0.322	0.254	0.179	0.098	0.021	C = 0.3 (Standard) Median Rank
30	0.256	0.191	0.131	0.072	0.015	C = 1 Median Rank
40	0.355	0.283	0.208	0.121	0.024	Unmodified
40	0.369	0.303	0.220	0.130	0.026	C = 0.3175 Median Rank
40	0.402	0.331	0.250	0.147	0.030	C = 0 Median Rank
40	0.366	0.298	0.215	0.128	0.026	C = 0.375 Median Rank
40	0.372	0.304	0.222	0.133	0.026	C = 0.3 (Standard) Median Rank
40	0.308	0.237	0.171	0.093	0.018	C = 1 Median Rank
50	0.408	0.336	0.253	0.137	0.038	Unmodified
50	0.428	0.353	0.266	0.149	0.042	C = 0.3175 Median Rank
50	0.461	0.382	0.294	0.167	0.047	C = 0 Median Rank
50	0.423	0.347	0.260	0.145	0.041	C = 0.375 Median Rank
50	0.430	0.355	0.269	0.150	0.042	C = 0.3 (Standard) Median Rank
50	0.357	0.288	0.207	0.108	0.027	C = 1 Median Rank

Table 4.21. CRAMER-VON MISES GOFT's Power Against EXPONENTIAL(1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.653	0.593	0.506	0.383	0.189	Unmodified
10	0.652	0.596	0.509	0.385	0.182	C = 0.3175 Median Rank
10	0.647	0.588	0.507	0.374	0.174	C = 0 Median Rank
10	0.655	0.595	0.507	0.383	0.183	C = 0.375 Median Rank
10	0.653	0.596	0.508	0.384	0.182	C = 0.3 (Standard) Median Rank
10	0.631	0.576	0.493	0.378	0.189	C = 1 Median Rank
20	0.903	0.873	0.818	0.733	0.519	Unmodified
20	0.901	0.870	0.819	0.726	0.513	C = 0.3175 Median Rank
20	0.897	0.868	0.815	0.718	0.503	C = 0 Median Rank
20	0.902	0.871	0.820	0.728	0.516	C = 0.375 Median Rank
20	0.901	0.870	0.819	0.726	0.514	C = 0.3 (Standard) Median Rank
20	0.897	0.865	0.813	0.729	0.515	C = 1 Median Rank
30	0.972	0.962	0.942	0.895	0.759	Unmodified
30	0.971	0.962	0.942	0.896	0.754	C = 0.3175 Median Rank
30	0.972	0.960	0.943	0.893	0.745	C = 0 Median Rank
30	0.971	0.962	0.942	0.897	0.759	C = 0.375 Median Rank
30	0.971	0.961	0.942	0.895	0.754	C = 0.3 (Standard) Median Rank
30	0.971	0.959	0.939	0.891	0.747	C = 1 Median Rank
40	0.993	0.989	0.984	0.964	0.889	Unmodified
40	0.993	0.989	0.985	0.965	0.887	C = 0.3175 Median Rank
40	0.993	0.990	0.985	0.965	0.883	C = 0 Median Rank
40	0.993	0.989	0.984	0.965	0.888	C = 0.375 Median Rank
40	0.993	0.990	0.985	0.965	0.887	C = 0.3 (Standard) Median Rank
40	0.992	0.989	0.982	0.962	0.886	C = 1 Median Rank
50	0.999	0.998	0.996	0.989	0.959	Unmodified
50	0.999	0.998	0.996	0.989	0.958	C = 0.3175 Median Rank
50	0.999	0.998	0.996	0.989	0.958	C = 0 Median Rank
50	0.999	0.998	0.996	0.989	0.958	C = 0.375 Median Rank
50	0.999	0.998	0.996	0.989	0.959	C = 0.3 (Standard) Median Rank
50	0.999	0.998	0.995	0.988	0.956	C = 1 Median Rank

Table 4.22. CRAMER-VON MISES GOFT's Power Against DBL EXPONENTIAL(2,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.708	0.655	0.577	0.467	0.254	Unmodified
10	0.707	0.656	0.582	0.469	0.251	C = 0.3175 Median Rank
10	0.702	0.651	0.574	0.458	0.240	C = 0 Median Rank
10	0.708	0.656	0.580	0.468	0.253	C = 0.375 Median Rank
10	0.708	0.656	0.582	0.468	0.252	C = 0.3 (Standard) Median Rank
10	0.690	0.637	0.565	0.463	0.256	C = 1 Median Rank
20	0.941	0.920	0.881	0.819	0.650	Unmodified
20	0.941	0.919	0.881	0.815	0.643	C = 0.3175 Median Rank
20	0.940	0.916	0.877	0.806	0.633	C = 0 Median Rank
20	0.941	0.920	0.880	0.817	0.646	C = 0.375 Median Rank
20	0.941	0.919	0.880	0.815	0.643	C = 0.3 (Standard) Median Rank
20	0.936	0.916	0.875	0.813	0.644	C = 1 Median Rank
30	0.989	0.981	0.973	0.947	0.864	Unmodified
30	0.989	0.981	0.973	0.947	0.861	C = 0.3175 Median Rank
30	0.989	0.981	0.972	0.946	0.854	C = 0 Median Rank
30	0.989	0.981	0.973	0.947	0.863	C = 0.375 Median Rank
30	0.989	0.981	0.973	0.946	0.861	C = 0.3 (Standard) Median Rank
30	0.987	0.982	0.971	0.946	0.859	C = 1 Median Rank
40	0.998	0.996	0.994	0.987	0.951	Unmodified
40	0.998	0.996	0.994	0.986	0.950	C = 0.3175 Median Rank
40	0.997	0.996	0.994	0.986	0.948	C = 0 Median Rank
40	0.998	0.996	0.994	0.987	0.950	C = 0.375 Median Rank
40	0.998	0.996	0.994	0.986	0.950	C = 0.3 (Standard) Median Rank
40	0.997	0.996	0.993	0.986	0.949	C = 1 Median Rank
50	1.000	1.000	0.999	0.997	0.989	Unmodified
50	1.000	1.000	0.999	0.997	0.989	C = 0.3175 Median Rank
50	1.000	1.000	0.999	0.997	0.988	C = 0 Median Rank
50	1.000	1.000	0.999	0.997	0.989	C = 0.375 Median Rank
50	1.000	1.000	0.999	0.997	0.989	C = 0.3 (Standard) Median Rank
50	1.000	1.000	1.000	0.997	0.987	C = 1 Median Rank

Table 4.23. CRAMER-VON MISES GOFT's Power Against WEIBULL(2,6) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.253	0.203	0.143	0.072	0.015	Unmodified
10	0.257	0.205	0.145	0.074	0.015	C = 0.3175 Median Rank
10	0.259	0.208	0.144	0.077	0.016	C = 0 Median Rank
10	0.256	0.204	0.145	0.073	0.016	C = 0.375 Median Rank
10	0.258	0.204	0.145	0.074	0.016	C = 0.3 (Standard) Median Rank
10	0.246	0.196	0.135	0.073	0.015	C = 1 Median Rank
20	0.339	0.274	0.196	0.115	0.034	Unmodified
20	0.338	0.278	0.199	0.113	0.033	C = 0.3175 Median Rank
20	0.351	0.278	0.202	0.113	0.032	C = 0 Median Rank
20	0.338	0.276	0.197	0.114	0.034	C = 0.375 Median Rank
20	0.340	0.277	0.200	0.113	0.033	C = 0.3 (Standard) Median Rank
20	0.331	0.264	0.191	0.112	0.034	C = 1 Median Rank
30	0.405	0.334	0.255	0.154	0.047	Unmodified
30	0.410	0.340	0.256	0.156	0.047	C = 0.3175 Median Rank
30	0.421	0.344	0.261	0.158	0.046	C = 0 Median Rank
30	0.409	0.337	0.256	0.156	0.048	C = 0.375 Median Rank
30	0.411	0.340	0.256	0.155	0.047	C = 0.3 (Standard) Median Rank
30	0.395	0.317	0.237	0.148	0.043	C = 1 Median Rank
40	0.468	0.394	0.316	0.211	0.066	Unmodified
40	0.474	0.402	0.322	0.212	0.066	C = 0.3175 Median Rank
40	0.484	0.409	0.324	0.210	0.065	C = 0 Median Rank
40	0.473	0.401	0.319	0.212	0.066	C = 0.375 Median Rank
40	0.475	0.402	0.322	0.213	0.066	C = 0.3 (Standard) Median Rank
40	0.453	0.383	0.305	0.202	0.063	C = 1 Median Rank
50	0.548	0.475	0.391	0.254	0.098	Unmodified
50	0.554	0.477	0.393	0.255	0.098	C = 0.3175 Median Rank
50	0.563	0.482	0.399	0.258	0.099	C = 0 Median Rank
50	0.554	0.477	0.391	0.255	0.099	C = 0.375 Median Rank
50	0.555	0.477	0.394	0.255	0.099	C = 0.3 (Standard) Median Rank
50	0.535	0.461	0.373	0.242	0.096	C = 1 Median Rank

Table 4.24. CRAMER-VON MISES GOFT's Power Against LOG-NORMAL(0,1) data; alpha = .20 ... .01

N	.20	.15	.10	.05	.01	PLOTTING POSITION
10	0.768	0.727	0.661	0.567	0.369	Unmodified
10	0.763	0.723	0.662	0.562	0.364	C = 0.3175 Median Rank
10	0.755	0.711	0.654	0.546	0.355	C = 0 Median Rank
10	0.766	0.724	0.662	0.563	0.367	C = 0.375 Median Rank
10	0.763	0.722	0.662	0.561	0.363	C = 0.3 (Standard) Median Rank
10	0.760	0.719	0.659	0.567	0.372	C = 1 Median Rank
20	0.963	0.949	0.925	0.883	0.780	Unmodified
20	0.963	0.948	0.924	0.881	0.775	C = 0.3175 Median Rank
20	0.963	0.946	0.922	0.877	0.764	C = 0 Median Rank
20	0.964	0.948	0.924	0.881	0.777	C = 0.375 Median Rank
20	0.963	0.948	0.924	0.881	0.776	C = 0.3 (Standard) Median Rank
20	0.960	0.946	0.922	0.884	0.779	C = 1 Median Rank
30	0.995	0.992	0.987	0.972	0.930	Unmodified
30	0.995	0.992	0.987	0.972	0.926	C = 0.3175 Median Rank
30	0.995	0.991	0.987	0.971	0.924	C = 0 Median Rank
30	0.995	0.992	0.987	0.973	0.929	C = 0.375 Median Rank
30	0.995	0.992	0.987	0.972	0.926	C = 0.3 (Standard) Median Rank
30	0.995	0.991	0.985	0.973	0.929	C = 1 Median Rank
40	0.999	0.999	0.998	0.996	0.982	Unmodified
40	0.999	0.999	0.998	0.996	0.981	C = 0.3175 Median Rank
40	0.999	0.999	0.999	0.996	0.981	C = 0 Median Rank
40	0.999	0.999	0.998	0.996	0.981	C = 0.375 Median Rank
40	0.999	0.999	0.998	0.996	0.981	C = 0.3 (Standard) Median Rank
40	0.999	0.999	0.998	0.996	0.981	C = 1 Median Rank
50	1.000	1.000	1.000	0.999	0.998	Unmodified
50	1.000	1.000	1.000	0.999	0.998	C = 0.3175 Median Rank
50	1.000	1.000	1.000	0.999	0.998	C = 0 Median Rank
50	1.000	1.000	1.000	0.999	0.998	C = 0.375 Median Rank
50	1.000	1.000	1.000	0.999	0.998	C = 0.3 (Standard) Median Rank
50	1.000	1.000	1.000	0.999	0.998	C = 1 Median Rank

V. Conclusions

This chapter is a presentation of the author's conclusions concerning this research thesis.

5.1 Review of Research Objectives

The purpose of this research has been to test five modifications to the A-D and the C-VM statistics to determine if an increase in the power of the goodness-of-fit test could be realized.

5.2 Conclusions

Two general trends are apparent from the power tables regardless of which table is being examined. The first is that as the sample size increases, so too increases the power. This makes sense because as more information is gained about a population, it becomes more apparent that it is not from the normal distribution. The second trend is that as α increases so does the power of that test. The reason for this is not quite as obvious, but to shed some light, consider again what α represents - the probability of wrongly rejecting the null hypothesis. When the analyst is less concerned about erroneously rejecting a true hypothesis, and he acts on this lessened concern by selecting a larger α, then it follows that more null hypotheses will be rejected, which increases the overall number of rejections, and subsequently, the power!

For discussion of the power studies, I will confine my analysis primarily to the $\alpha = .05$ levels of confidence since a good number of hypothesis tests use this confidence level. One side note concerning the validity of the FORTRAN subroutine POWERFUL which performs the power studies and generates the tables: when running the program against data from normal(0,1) data, as expected, the power of both the A-D and C-VM tests is equal to α; .20, .15, .10, .05, and .01. If this

were not the case, then something would be wrong because α is the probability of rejecting the null hypothesis when in fact it is true.

However, when testing against the uniform(0,1) distribution , both the A-D and the C-VM tests are fairly powerful with the A-D being the stronger of the two. The strongest individual test is the modified A-D test, c=0, n=50, where the power is .574. At all α and n levels, the modified A-D and the modified C-VM where c=0 are the most powerful. All modifications to the A-D and the C-VM statistics, except for c=1, show higher power than the tests with the unmodified statistics.

It is a different story for the test powers against the beta(3,2) distribution. Neither the A-D nor the C-VM test are very powerful. The strongest individual test is the modified C-VM test, c=0, n=50, where the power is .167. Again, as with the uniform distribution, at all α and n levels, the modified A-D and the modified C-VM where c=0 are the most powerful. All modifications to the A-D statistic, except for c=1, show at least as high a power, if not higher than the tests with the unmodified A D statistic. All modifications to the C-VM statistic, except for c=1, show higher power than the tests with the unmodified C-VM statistic.

Against the exponential(1) distribution, both the A-D and the C-VM tests are very powerful with the A-D being ever-so-slightly the stronger of the two. There is a 6-way tie for the strongest individual test. It is amongst all of the A-D tests at n=50 where the power is .995. Unlike the two previous distributions, there is no pattern for any one statistic dominating the power studies at all α and n levels, parity amongst the statistics seems to be the key word for the exponential(1). All modifications to the A-D and the C-VM statistics show approximately the same power as the tests with the unmodified statistics.

Almost identical to the power against the exponential(1) distribution, for the double exponential(2,1,) both the A-D and the C-VM tests are very powerful with the A-D tests being, again, slightly stronger than the C-VM tests. There is a 6-way tie for the strongest individual test. It was amongst all of the A-D tests at n=50

where the power is .999. Also like the exponential(1), there is no apparent pattern of any one statistic dominating the power studies at all α and n levels. All modifications to the A-D and the C-VM statistics show approximately the same power as the tests with the unmodif tatistics.

The po dy against the Weibull(2,6) distribution, shows both the A-D and the C-VM tests having low power. For the small sample sizes, there is a rough similarity between the powers of the A-D and C-VM tests, but as the sample sizes approach 50, the A-D test is clearly the stronger. There is a 2-way tie for the strongest individual test between the unmodified A-D test and the modified A-D c=0 test, at n=50, where the power is .305. There is not a pattern of any one statistic dominating the power studies at all α and n levels. All modifications to the A-D and the C-VM statistics, except for c=1, show at least as high a power as the tests with the unmodified statistics.

Against the last distribution tested, the log-normal(0,1), both the A-D and the C-VM tests are very powerful. As with the majority of the other distributions, the A-D test is slightly more powerful. There is again a 6-way tie for the strongest individual test. It is amongst all of the A-D tests at n=50 where the power is .999. And there is no pattern of any one statistic dominating the power studies at all α and n levels. All modifications to the A-D and the C-VM statistics show approximately the same power as the tests with the unmodified statistics.

To summarize, the results of the power studies against the exponential, double exponential, Weibull and log-normal show little or no improvements over the power of the unmodified statistics via eight of the modifications; in fact, two of the new statistics show a marked decrease in power against these distributions. In the case of the uniform and beta distributions, the power increase is more pronounced with the c=0 modification to the A-D and C-VM statistic with the A-D statistic being the stronger (Overall, there is a trend of the A-D tests being the more powerful versus the C-VM tests). Because at the very least, no power is lost against any of

these distributions, but actually significantly gained in two cases, this research has shown that a more powerful goodness-of-fit test for normality with the parameters estimated from the sample is available in the modified A-D statistic defined by

$$AD_{MEDIAN,C=0} = -n - 2\sum_{i=1}^{n} \frac{i}{n+1}(\log_e(U_i) + log_e(1 - U_{n-i+1})).$$

Appendix A. FORTRAN Programs

```
******************************************************************
*                                                                *
*      THIS PROGRAM DETERMINES THE CRITICAL VALUES WITH          *
*   THE MODIFIED ANDERSON-DARLING & CRAMER-VON MISES STATISTICS  *
*         VALUES FOR THE NORMAL DISTRIBUTION.                    *
*                                                                *
*      THIS PROGRAM ALSO GENERATES THE POWER STUDIES.            *
*                                                                *
*      MANY WRITE STATEMENTS HAVE BEEN COMMENTED OUT IN THIS,    *
*   CODE. THE PURPOSE OF THESE WRITE STATEMENTS WAS FOR VERIF-   *
*   ICATION AND TESTING PURPOSES AND THEY HAVVE BEEN LEFT IN     *
*   FOR FUTURE USERS BENEFIT.                                    *
*                                                                *
*   MANY THANKS TO KAHYA AND REAM FOR THE BULK OF THIS CODE.     *
*   WITHOUT THEIR EFFORTS, MY THESIS WOULD NOT HAVE GONE AS FAR  *
*   AS IT HAS.                                                   *
*                                                                *
******************************************************************


C     ** SOME COMMENTS ABOUT THE VARIABLES USED **
C
C     S = # OF REPETITIONS, 10,000 THIS THESIS
C     N = SAMPLE SIZE, 4, 5, ..., 50
C     ANDER(10003) = AD STATISTIC VALUE ARRAY
C     SEED = INITIAL VALUE NEEDED TO START RANDOM NUM GENS
C          THROUGH IMSL ROUTINE RNSET
C     "XX(*) = OUTPUT VECTOR FOR NORMAL DEVIATES

      INTEGER S,J,N,K,SS,SEED,III,I,RUNTYPE,PASSNO,IJ
      REAL XX(50),ANDER(0:10003),DARL
      REAL CRIT80,CRIT85,CRIT90,CRIT95,CRIT99,ANORDF
      CHARACTER*32 COMBONAM
      EXTERNAL RNNOR,RNSET,SVRGN,ANORDF

C     THE NUMBER OF REPETITIONS IS
      S = 10000

      PRINT *
      WRITE(*,1)
```

```
1     FORMAT('THIS PROGRAM WILL GEN THE AD AND CVM CRIT VALS FOR ONE',
     +' OF THE SEVEN CHOICES BELOW (N = 4 TO 50) FOR THE N(0,1) DISTR ',
     +'AND THEN PREPARE A POWER STUDY USING DATA FROM THE UNIF, BETA ',
     +' EXP, DBL EXP, WEIB AND LNORM DISTS (N = 10,50,10).'//,
     +'YOU MUST CHOOSE FROM THE FOLLOWING OPTS:',//,
     +'1 - USE UNMODIFIED PLOTTING POSITION WITH AD & CVM STATS',/,
     +'2 - USE KAHYA'S MESSED UP MEDIAN RANK PLOTTING POSITION ',
     +'WITH AD & CVM STATS',/,
     +'3 - USE C = 0.3175 MEDIAN RANK PLOTTING POSITION ',
     +'WITH AD & CVM STATS',/,
     +'4 - USE C = 0 MEDIAN (KAHYA'S MEAN) RANK PLOTTING POSITION',
     +'WITH AD & CVM STATS',/,
     +'5 - USE C = 0.375 MEDIAN RANK PLOTTING POSITION WITH AD & CVM ',
     +'STATS',/,
     +'6 - USE C = 0.3 (STANDARD) MEDIAN RANK PLOTTING POSITION WITH ',
     +'AD & CVM STATS',/,
     +'7 - USE C = 1 MEDIAN RANK PLOTTING POSITION WITH AD & CVM STATS',
     +//,'ENTER YOUR SELECTION AS I.E. 3 OR 1.'//)

      READ(*,2)RUNTYPE
2     FORMAT(I1)

      IF(RUNTYPE.EQ.1)
     +    COMBONAM='UNMODIFIED                      '
      IF(RUNTYPE.EQ.2)
     +    COMBONAM='KAHYA'S MESSED UP MEDIAN RANK '
      IF(RUNTYPE.EQ.3)
     +    COMBONAM='C = 0.3175 MEDIAN RANK          '
      IF(RUNTYPE.EQ.4)
     + COMBONAM='C = 0 MEDIAN RANK                   '
      IF(RUNTYPE.EQ.5)
     +    COMBONAM='C = 0.375 MEDIAN RANK           '
      IF(RUNTYPE.EQ.6)
     +    COMBONAM='C = 0.3 (STANDARD) MEDIAN RANK'
      IF(RUNTYPE.EQ.7)
     +    COMBONAM='C = 1 MEDIAN RANK               '

      IF(RUNTYPE.LT.1.OR.RUNTYPE.GT.7) STOP

      IF(RUNTYPE.EQ.1)THEN
        OPEN(UNIT=1,FILE='adcv1.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv1.dat',STATUS='UNKNOWN')
```

```
      ENDIF
      IF(RUNTYPE.EQ.2)THEN
        OPEN(UNIT=1,FILE='adcv2.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv2.dat',STATUS='UNKNOWN')
      ENDIF
      IF(RUNTYPE.EQ.3)THEN
        OPEN(UNIT=1,FILE='adcv3.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv3.dat',STATUS='UNKNOWN')
      ENDIF
      IF(RUNTYPE.EQ.4)THEN
        OPEN(UNIT=1,FILE='adcv4.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv4.dat',STATUS='UNKNOWN')
      ENDIF
      IF(RUNTYPE.EQ.5)THEN
        OPEN(UNIT=1,FILE='adcv5.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv5.dat',STATUS='UNKNOWN')
      ENDIF
      IF(RUNTYPE.EQ.6)THEN
        OPEN(UNIT=1,FILE='adcv6.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv6.dat',STATUS='UNKNOWN')
      ENDIF
      IF(RUNTYPE.EQ.7)THEN
        OPEN(UNIT=1,FILE='adcv7.dat',STATUS='UNKNOWN')
        OPEN(UNIT=2,FILE='cvmcv7.dat',STATUS='UNKNOWN')
      ENDIF

      DO 3000 PASSNO = 1,2

        IF(PASSNO.EQ.1)THEN
          SEED = 402958961
        ELSE
          SEED = 510496597
        ENDIF
        CALL RNSET(SEED)

        DO 2000 N = 4,50

**************************************************************
* INITIALIZE EVERYTHING FOR THIS RUN OF SAMPLE SIZE N.        *
**************************************************************

          J=0
```

```
      K=0
      SS=0
      III=0
      I=0
      DO 5 I = 1,N
        XX(I)=0.0
5     CONTINUE
      DO 6 I = 0,S+2
        ANDER(I)=0.0
6     CONTINUE
      DARL=0.0
      CRIT80=0.0
      CRIT85=0.0
      CRIT90=0.0
      CRIT95=0.0
      CRIT99=0.0

************************************************
* DO THE GRUNT WORK FOR THIS SAMPLE OF SIZE N.*
************************************************

      DO 11 J = 1,S
C
C       CALL RNNOR TO GENERATE SET OF ORDER STATISTICS
C       FROM THE NORMAL DISTRIBUTION
C
        CALL RNNOR(N,XX)

C       CALL STANDAR TO STANDARDIZE THE RANDOM DEVIATES.
C       THE FOLLOWING COMPUTATIONS WILL
C       USE THESE STANDARDIZED VALUES.

C       WRITE(*,*)'NORMAL DEVIATES   = ',(XX(IJ),IJ=1,4)
        CALL STANDAR(XX,N)
C       WRITE(*,*)'STANDARDIZED DATA = ',(XX(IJ),IJ=1,4)

C       CALL SVRGN TO SORT THE DATA INTO ASCENDING ORDER
C
        CALL SVRGN(N,XX,XX)
C       WRITE(*,*)'SORTED DATA       = ',(XX(IJ),IJ=1,4)

        DO 22 K=1,N
```

```
C           USE ANORDF TO DTRMN THE CDF VALUES OF THE NORMAL
C           DISTRIBUTION FOR THE STANDARDIZED DATA.

            XX(K) = ANORDF(XX(K))
22        CONTINUE

C         WRITE(*,*)'CDFED DATA (Uis)  = ',(XX(IJ),IJ=1,4)


C         CALL ANDERSON/CRAMER TO FIND THE AD STAT VALUES

          IF(PASSNO.EQ.1)CALL ANDERSON(N,XX,DARL,RUNTYPE)
          IF(PASSNO.EQ.2)CALL CRAMER(N,XX,DARL,RUNTYPE)

          ANDER(J) = DARL
C         WRITE(*,*)'THIS AD STAT IS ',ANDER(J)

11      CONTINUE

        ANDER(0) = 0.0
        SS = S+1

C       CALL SVRGN TO SORT THE STATS TO DTRMN THE PERCENTILES
C       OF THE ARRAY IN ORDER TO FIND THE CRITICAL VALUES.

        CALL SVRGN(SS,ANDER,ANDER)

C       CALL EXTRA TO XTRPLT THE DATA FOR THE POINTS AT "0,"
C       AND "N+1."  OUTPUT VECTOR IS AN ARRAY OF LENGTH "N+2."

        CALL EXTRA(S,ANDER)

C       CALL VALUES TO DETERMINE THE CRITICAL VALUES

        CALL VALUES(ANDER,CRIT80,CRIT85,CRIT90,CRIT95,CRIT99,SS)
        WRITE(PASSNO,1599) N,CRIT80,CRIT85,CRIT90,CRIT95,CRIT99
1599    FORMAT(1X,I2,5(2X,F6.4))

***********************************

2000  CONTINUE
```

```
 3000  CONTINUE

       DO 3010 I = 1,2
          REWIND(I)
 3010  CONTINUE

       PRINT *
       PRINT *,'  The critical value tables for the menu ',
      +'option you selected are in adcv(opt).dat & cvmcv(opt).dat. '
       PRINT *

       CALL POWERFUL(COMBONAM,RUNTYPE)

       PRINT *,'The Anderson-Darling power studies for the uniform, ',
      +'beta, exponential, double exponential weibull, and lognormal ',
      +'are in these files: adunif(opt).dat, adbeta(opt).dat, ',
      +'adexpn(opt).dat, ',
      +'addblx(opt).dat, adweib(opt).dat, and adlogn(opt).dat.'
       PRINT *
       PRINT *,'The Cramer-von Mises power studies for the uniform, ',
      +'beta, exponential, double exponential weibull, and lognormal ',
      +'are in these files: cvmunif(opt).dat, cvmbeta(opt).dat, ',
      +'cvmexpn(opt).dat, ',
      +'cvmdblx(opt).dat, cvmweib(opt).dat, and cvmlogn(opt).dat.'
       PRINT *

       STOP
       END

*****************************************************************
*                                                               *
*     USING THE TECHNIQUE EXPLAINED IN CHAPTER 3, FIND THE      *
*     CRITICAL VALUES                                           *
*                                                               *
*****************************************************************

       SUBROUTINE CAN(Y1,Y2,D1,D2,Y,RES)
       REAL M,B,Y1,Y2,D1,D2,Y,RES
       IF((D2-D1).EQ.0.0)D2 = D2 * 1.00001
       M = (Y2-Y1)/(D2-D1)
```

```
      B = Y1 - M*D1
      RES = (Y-B)/M
      RETURN
      END

*****************************************************************
*                                                               *
* THE FOLLOWING SUB CALCS THE AD STAT VALUES; DEPENDING ON      *
* THE VALUE OF RUNTYPE, THE APPROPRIATE MOD IS USED (1:7).      *
*                                                               *
*****************************************************************

      SUBROUTINE ANDERSON(K,R,ANDER,RUNTYPE)
      INTEGER K,I,RUNTYPE
      REAL R(50),TOTAL,XX,YY,ANDER,ZZ
      TOTAL = 0.0
c      print *
c      WRITE(*,100)
c      print *
c 100   FORMAT('N  I  I-1      Ui    ln(Ui)  1-Un-i+1    ln(1-Un-i+1)  '
c     +' ln(Ui)+ln(',
c     +'1-Un-i+1)  (I-1)*(ln(Ui)+ln(1-Un-i+1))  SUM'/,
c     +       '     ---                                                ',
c     +'           ',
c     +'          -----                               '/,
c     +       '     N-1                                                ',
c     +'                    N-1')

     DO 11 I = 1,K
       IF(R(I).LE.0.0) R(I) = 0.0001
       ZZ = 1.0 - R(K-I+1)
       IF(ZZ.LE.0.0) ZZ = 0.0001
       XX = LOG(R(I))
       YY = LOG(ZZ)
c                                              UNMODIFIED PLOTTING POSITION
       IF(RUNTYPE.EQ.1)
     +    TOTAL = (2.0*REAL(I)-1.0)*(XX+YY)/(2.0*REAL(K)) + TOTAL

c                         KAHYA'S MESSED UP MEDIAN RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.2)
     +   TOTAL = (REAL(I)-0.3175)*(XX+YY)/(REAL(K)-0.365) + TOTAL
```

```
c                             KAHYA'S PROPER MEDIAN RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.3)
     +   TOTAL = (REAL(I)-0.3175)*(XX+YY)/(REAL(K)+0.365) + TOTAL

c                       C = 0 MEDIAN (KAHYA'S MEAN) RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.4)
     +   TOTAL = (REAL(I)*(XX+YY))/(REAL(K)+1.0) + TOTAL

c                                  GWINN'S MEDIAN RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.5)
     +   TOTAL = (REAL(I)-0.375)*(XX+YY)/(REAL(K)+0.25) + TOTAL

c                                STANDARD MEDIAN RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.6)
     +   TOTAL = (REAL(I)-0.3)*(XX+YY)/(REAL(K)+0.4) + TOTAL

c                                   C = 1 MEDIAN RANK PLOTTING POSITION
       IF(RUNTYPE.EQ.7)
     +   TOTAL = (REAL(I)-1.0)*(XX+YY)/(REAL(K)-1.0) + TOTAL
c       WRITE(*,150)K,I,(REAL(I)-1.0)/(REAL(K)-1.0),R(I),XX,ZZ,YY,XX+YY,
c    +  (REAL(I)-1.0)*(XX+YY)/(REAL(K)-1.0),TOTAL
c 150    FORMAT(I1,2X,I1,2X,4(F7.4,1X),2X,F7.4,8X,F7.4,15X,F7.4,20X,F7.4)

11    CONTINUE
c      print *
c      print *,'AD = -N -2. * TOTAL'
c      PRINT *
      ANDER = -REAL(K) - 2.0 * TOTAL
      RETURN

      END

***************************************************************
*                                                             *
* THE FOLLOWING SUBR CALCS THE CVM STAT VALUES; DEPENDING ON  *
* THE VALUE OF RUNTYPE, THE APPROPRIATE MOD IS USED (1:5).    *
*                                                             *
***************************************************************

      SUBROUTINE CRAMER(N,R,ANDER,RUNTYPE)
      INTEGER I,N,RUNTYPE
      REAL R(50),TOTAL,ANDER,MEDIAN
```

```
      TOTAL = 0.0
      DO 22 I = 1,N
c                                        UNMODIFIED PLOTTING POSITION
        IF(RUNTYPE.EQ.1)
     +     MEDIAN = (2.0*REAL(I)-1.0)/(2.0*REAL(N))

c                     KAHYA'S MESSED UP MEAN RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.2)
     +     MEDIAN = (REAL(I)-0.3175)/(REAL(N)-0.365)

c                        KAHYA'S PROPER MEAN RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.3)
     +     MEDIAN = (REAL(I)-0.3175)/(REAL(N)+0.365)

c                C = 0 MEDIAN (KAHYA'S MEAN) RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.4)
     +     MEDIAN = REAL(I)/(REAL(N)+1.0)

c                                GWINN'S MEDIAN RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.5)
     +     MEDIAN = (REAL(I)-0.375)/(REAL(N)+0.25)

c                              STANDARD MEDIAN RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.6)
     +     MEDIAN = (REAL(I)-0.3)/(REAL(N)+0.4)

c                                   C = 1 MEDIAN RANK PLOTTING POSITION
        IF(RUNTYPE.EQ.7)
     +     MEDIAN = (REAL(I)-1.0)/(REAL(N)-1.0)

        TOTAL = TOTAL + (R(I) - MEDIAN)**2
22    CONTINUE
      ANDER = (1.0/(12.0*REAL(N))) + TOTAL
      RETURN
      END

**************************************************************
*                                                            *
*     THIS SUBROUTINE STANDARDIZES ALL THE XX(I) DATA        *
*                                                            *
**************************************************************
```

```
      SUBROUTINE STANDAR(X,N)
      INTEGER I,N
      REAL X(50),XSUM,XBAR S,XOUT
      XSUM = 0.0
      XOUT = 0.0
      DO 100 I = 1,N
        XSUM = XSUM + X(I)
100   CONTINUE
      XBAR = XSUM/N
      DO 200 I = 1,N
        XOUT = XOUT + (X(I) - XBAR)**2
200   CONTINUE
      S = SQRT(XOUT/(N - 1))
      DO 300 I = 1,N
        X(I) = (X(I) - XBAR)/S
300   CONTINUE
      RETURN
      END

*************************************************************
*                                                           *
*     THIS SUBROUTINE EXTRAPOLATES THE ANDER(I) DATA        *
*     TO GENERATE ANDER(0) AND ANDER(S+1) FOR COMPUTATION   *
*     OF THE FIVE CRITICAL VALUES.                          *
*                                                           *
*************************************************************

      SUBROUTINE EXTRA(N,D)
      INTEGER N,N0,N1
      REAL Y1,Y2,D(0:10003),D1,D2,ZZ
      Y1 = 0.5/N
      Y2 = 1.5/N
      D1 = D(1)
      D2 = D(2)
      CALL CAN(Y1,Y2,D1,D2,0.0,ZZ)
      IF(ZZ.GE.0.0) THEN
        D(0) = ZZ
      ELSE
        D(0) = 0.0
      ENDIF
      Y1 = (REAL(N) - 1.5)/N
      Y2 = (REAL(N) - 0.5)/N
```

```
      NO = N-1
      D1 = D(NO)
      D2 = D(N)
      CALL CAN(Y1,Y2,D1,D2,1.0,ZZ)
      N1 = N + 1
      D(N1) = ZZ
      RETURN
      END


*****************************************************************
*                                                               *
*     THE FOLLOWING SUB DETERMINES THE %TILES AND FINDS         *
*     THE CRITICAL VALUES BY EVOKING THE SUB CAN                *
*                                                               *
*****************************************************************

      SUBROUTINE VALUES(D,CRIT80,CRIT85,CRIT90,CRIT95,CRIT99,N)
      INTEGER I,N,NN
      REAL D(0:10003),Y(0:10003),C80,C90,C95,C99,C85,
     + Y79,D79,Y81,D81,DIF90,Y89,Y91,D89,D91,DIF95,DIF80,
     + Y94,Y96,D94,D96,DIF99,Y98,Y100,D98,D100,DIF85,
     + Y84,D84,Y86,D86,CRIT85,CRIT80,CRIT90,CRIT95,CRIT99
      DO 100 I = 1,N
        Y(I) = (REAL(I) - 0.5)/REAL(N)
100   CONTINUE
      Y(0) = 0.0
      NN = N + 1
      Y(NN) = 1.0
      C80 = 1000.0
      C85 = 1000.0
      C90 = 1000.0
      C95 = 1000.0
      C99 = 1000.0
      DO 200 I = NN,0,-1
        IF (Y(I).LE.0.75) GO TO 300
        IF (Y(I).GT.0.75.AND.Y(I).LE.0.80) THEN

C         GET THE DESIRED %TILE AT 80%
          DIF80 = .80 - Y(I)
          IF (DIF80.LE.C80) THEN
            C80 = DIF80
```

```
            Y79 = Y(I)
            D79 = D(I)
            Y81 = Y(I+1)
            D81 = D(I+1)
          ENDIF

        ELSEIF (Y(I).GT.0.80.AND.Y(I).LE.0.85) THEN

C         GET THE DESIRED %TILE AT 85%
          DIF85 = .85 - Y(I)
          IF (DIF85.LE.C85) THEN
            C85 = DIF85
            Y84 = Y(I)
            D84 = D(I)
            Y86 = Y(I+1)
            D86 = D(I+1)
          ENDIF

        ELSEIF (Y(I).GT.0.85.AND.Y(I).LE.0.90) THEN

C         GET THE DESIRED %TILE AT 90%
          DIF90 = .90 - Y(I)
          IF (DIF90.LE.C90) THEN
            C90 = DIF90
            Y89 = Y(I)
            D89 = D(I)
            Y91 = Y(I+1)
            D91 = D(I+1)
          ENDIF

        ELSEIF (Y(I).GT.0.90.AND.Y(I).LE.0.95) THEN

C         GET THE DESIRED %TILE AT 95%
          DIF95 = .95 - Y(I)
          IF (DIF95.LE.C95) THEN
            C95 = DIF95
            Y94 = Y(I)
            D94 = D(I)
            Y96 = Y(I+1)
            D96 = D(I+1)
          ENDIF
```

```
        ELSEIF (Y(I).GT.0.95.AND.Y(I).LE.0.99) THEN
C         GET THE DESIRED %TILE AT 99%
          DIF99 = .99 - Y(I)
          IF (DIF99.LE.C99) THEN
            C99 = DIF99
            Y98 = Y(I)
            D98 = D(I)
            Y100 = Y(I+1)
            D100 = D(I+1)
          ENDIF

        ENDIF
200   CONTINUE

300   IF (DIF80.EQ.0.0) THEN
        CRIT80 = D79
      ELSE
C       COMPUTE THE CRIT VALUE AT SIGNIFICANCE LEVEL = .20
        CALL CAN(Y79,Y81,D79,D81,.80,CRIT80)
      ENDIF

      IF (DIF85.EQ.0.0) THEN
        CRIT85 = D84
      ELSE
C       COMPUTE THE CRIT VALUE AT SIGNIFICANCE LEVEL = [illegible]
        CALL CAN(Y84,Y86,D84,D86,.85,CRIT85)
      ENDIF

      IF (DIF90.EQ.0.0) THEN
        CRIT90 = D89
      ELSE
C       COMPUTE THE CRIT VALUE AT SIGNIFICANCE LEVEL = .10
        CALL CAN(Y89,Y91,D89,D91,.90,CRIT90)
      ENDIF

      IF (DIF95.EQ.0.0) THEN
        CRIT95 = D94
      ELSE
C       COMPUTE THE CRIT VALUE AT SIGNIFICANCE LEVEL = .05
        CALL CAN(Y94,Y96,D94,D96,.95,CRIT95)
      ENDIF
```

```
      IF (DIF99.EQ.0.0) THEN
        CRIT99 = D98
      ELSE
C       COMPUTE THE CRIT VALUE AT SIGNIFICANCE LEVEL = .01
        CALL CAN(Y98,Y100,D98,D100,.99,CRIT99)
      ENDIF
      RETURN
      END


*********************************************************************
*                                                                   *
*          THE POWERFUL SUB EXECUTES THE POWER STUDY                *
*       FOR THE ANDERSON-DARLING AND THE CRAMER-VON MISES           *
*       TESTS WITH THE SIX DISTRIBUTIONS AT DIFFERENT ALPHA         *
*       LEVELS, .20, .15, .10, .05, .01, AND AT DIFFERENT SAMPLE    *
*       SIZES, 10, 20, 30, 40, 50                                   *
*                                                                   *
*********************************************************************

C     ***** SOME COMMENTS ABOUT THE VARIABLES USED *****

C     POWER(*) = THE REJECTION %AGE AT THE GIVEN LEVEL

      SUBROUTINE POWERFUL(COMBONAM,RUNTYPE)
      INTEGER I,J,L,N,MAX,NUMDIST,RUNTYPE,IREPSIZE
      INTEGER DISTRIB,STATTYPE,ALPHA,IREC
      INTEGER SEED(6)
      REAL S(50),CRITVAL(2,5,5),DUMMY(6)
      REAL Y,P,ANDAR,AND,CRAMER1,CRAMER2,POWER(6,2,5),ANORDF
      CHARACTER*20 DISTNAM(6)
      CHARACTER*32 COMBONAM
      CHARACTER*34 STATNAM(2)
      EXTERNAL RNUN,RNBET,RNEXP,RNEXT,RNWIB,SSCAL,RNLNL,
     +         SVRGN,ANORDF
      DATA DISTNAM/'UNIFORM(0,1)','BETA(3,2)','EXPONENTIAL(1)'
     +     ,'DBL EXPONENTIAL(2,1)','WEIBULL(2,6)','LOGNORMAL(0,1)'/
      DATA STATNAM/'ANDERSON-DARLING                  ',
     +            'CRAMER-VON MISES                  '/
      DATA SEED/77757985,172278979,536814143,751957587,4638907,
     +         271411103/
```

```
C                                       READ 1ST RECORD FROM EA FILE,
C                                       THIS CONTAINS CRITVALS FOR N=4
      READ(1,*)(DUMMY(I),I=1,6)
      READ(2,*)(DUMMY(I),I=1,6)

      DO 10 STATTYPE = 1,2
         DO 10 N = 10,50,10
C                                 SKIP NEXT FIVE RECORDS
            DO 2 IREC = 1,5
               READ(STATTYPE,*)(DUMMY(I),I=1,6)
2           CONTINUE

            READ(STATTYPE,*)DUMMY(1),(CRITVAL(STATTYPE,ALPHA,N/10),
     +                           ALPHA=1,5)
C                                SKIP NEXT FOUR RECORDS
            IF(N.LT.50)THEN
              DO 7 IREC = 1,4
                 READ(STATTYPE,*)(DUMMY(I),I=1,6)
7             CONTINUE
            ENDIF
10    CONTINUE

      IREPSIZE = 5000

      NUMDIST = 6

      DO 20 I = 1,2
        CLOSE(I)
20    CONTINUE

      MAX = 50

      IF(RUNTYPE.EQ.1)THEN
        OPEN(UNIT=10,FILE='adunif1.dat',STATUS='unknown')
        OPEN(UNIT=11,FILE='adbeta1.dat',STATUS='unknown')
        OPEN(UNIT=12,FILE='adexpn1.dat',STATUS='unknown')
        OPEN(UNIT=13,FILE='addblx1.dat',STATUS='unknown')
        OPEN(UNIT=14,FILE='adweib1.dat',STATUS='unknown')
        OPEN(UNIT=15,FILE='adlogn1.dat',STATUS='unknown')
        OPEN(UNIT=20,FILE='cvmunif1.dat',STATUS='unknown')
        OPEN(UNIT=21,FILE='cvmbeta1.dat',STATUS='unknown')
```

```
   OPEN(UNIT=22,FILE='cvmexpn1.dat',STATUS='unknown')
   OPEN(UNIT=23,FILE='cvmdblx1.dat',STATUS='unknown')
   OPEN(UNIT=24,FILE='cvmweib1.dat',STATUS='unknown')
   OPEN(UNIT=25,FILE='cvmlogn1.dat',STATUS='unknown')
 ENDIF
 IF(RUNTYPE.EQ.2)THEN
   OPEN(UNIT=10,FILE='adunif2.dat',STATUS='unknown')
   OPEN(UNIT=11,FILE='adbeta2.dat',STATUS='unknown')
   OPEN(UNIT=12,FILE='adexpn2.dat',STATUS='unknown')
   OPEN(UNIT=13,FILE='addblx2.dat',STATUS='unknown')
   OPEN(UNIT=14,FILE='adweib2.dat',STATUS='unknown')
   OPEN(UNIT=15,FILE='adlogn2.dat',STATUS='unknown')
   OPEN(UNIT=20,FILE='cvmunif2.dat',STATUS='unknown')
   OPEN(UNIT=21,FILE='cvmbeta2.dat',STATUS='unknown')
   OPEN(UNIT=22,FILE='cvmexpn2.dat',STATUS='unknown')
   OPEN(UNIT=23,FILE='cvmdblx2.dat',STATUS='unknown')
   OPEN(UNIT=24,FILE='cvmweib2.dat',STATUS='unknown')
   OPEN(UNIT=25,FILE='cvmlogn2.dat',STATUS='unknown')
 ENDIF
 IF(RUNTYPE.EQ.3)THEN
   OPEN(UNIT=10,FILE='adunif3.dat',STATUS='unknown')
   OPEN(UNIT=11,FILE='adbeta3.dat',STATUS='unknown')
   OPEN(UNIT=12,FILE='adexpn3.dat',STATUS='unknown')
   OPEN(UNIT=13,FILE='addblx3.dat',STATUS='unknown')
   OPEN(UNIT=14,FILE='adweib3.dat',STATUS='unknown')
   OPEN(UNIT=15,FILE='adlogn3.dat',STATUS='unknown')
   OPEN(UNIT=20,FILE='cvmunif3.dat',STATUS='unknown')
   OPEN(UNIT=21,FILE='cvmbeta3.dat',STATUS='unknown')
   OPEN(UNIT=22,FILE='cvmexpn3.dat',STATUS='unknown')
   OPEN(UNIT=23,FILE='cvmdblx3.dat',STATUS='unknown')
   OPEN(UNIT=24,FILE='cvmweib3.dat',STATUS='unknown')
   OPEN(UNIT=25,FILE='cvmlogn3.dat',STATUS='unknown')
 ENDIF
 IF(RUNTYPE.EQ.4)THEN
   OPEN(UNIT=10,FILE='adunif4.dat',STATUS='unknown')
   OPEN(UNIT=11,FILE='adbeta4.dat',STATUS='unknown')
   OPEN(UNIT=12,FILE='adexpn4.dat',STATUS='unknown')
   OPEN(UNIT=13,FILE='addblx4.dat',STATUS='unknown')
   OPEN(UNIT=14,FILE='adweib4.dat',STATUS='unknown')
   OPEN(UNIT=15,FILE='adlogn4.dat',STATUS='unknown')
   OPEN(UNIT=20,FILE='cvmunif4.dat',STATUS='unknown')
   OPEN(UNIT=21,FILE='cvmbeta4.dat',STATUS='unknown')
```

```
    OPEN(UNIT=22,FILE='cvmexpn4.dat',STATUS='unknown')
    OPEN(UNIT=23,FILE='cvmdblx4.dat',STATUS='unknown')
    OPEN(UNIT=24,FILE='cvmweib4.dat',STATUS='unknown')
    OPEN(UNIT=25,FILE='cvmlogn4.dat',STATUS='unknown')
  ENDIF
  IF(RUNTYPE.EQ.5)THEN
    OPEN(UNIT=10,FILE='adunif5.dat',STATUS='unknown')
    OPEN(UNIT=11,FILE='adbeta5.dat',STATUS='unknown')
    OPEN(UNIT=12,FILE='adexpn5.dat',STATUS='unknown')
    OPEN(UNIT=13,FILE='addblx5.dat',STATUS='unknown')
    OPEN(UNIT=14,FILE='adweib5.dat',STATUS='unknown')
    OPEN(UNIT=15,FILE='adlogn5.dat',STATUS='unknown')
    OPEN(UNIT=20,FILE='cvmunif5.dat',STATUS='unknown')
    OPEN(UNIT=21,FILE='cvmbeta5.dat',STATUS='unknown')
    OPEN(UNIT=22,FILE='cvmexpn5.dat',STATUS='unknown')
    OPEN(UNIT=23,FILE='cvmdblx5.dat',STATUS='unknown')
    OPEN(UNIT=24,FILE='cvmweib5.dat',STATUS='unknown')
    OPEN(UNIT=25,FILE='cvmlogn5.dat',STATUS='unknown')
  ENDIF
  IF(RUNTYPE.EQ.6)THEN
    OPEN(UNIT=10,FILE='adunif6.dat',STATUS='unknown')
    OPEN(UNIT=11,FILE='adbeta6.dat',STATUS='unknown')
    OPEN(UNIT=12,FILE='adexpn6.dat',STATUS='unknown')
    OPEN(UNIT=13,FILE='addblx6.dat',STATUS='unknown')
    OPEN(UNIT=14,FILE='adweib6.dat',STATUS='unknown')
    OPEN(UNIT=15,FILE='adlogn6.dat',STATUS='unknown')
    OPEN(UNIT=20,FILE='cvmunif6.dat',STATUS='unknown')
    OPEN(UNIT=21,FILE='cvmbeta6.dat',STATUS='unknown')
    OPEN(UNIT=22,FILE='cvmexpn6.dat',STATUS='unknown')
    OPEN(UNIT=23,FILE='cvmdblx6.dat',STATUS='unknown')
    OPEN(UNIT=24,FILE='cvmweib6.dat',STATUS='unknown')
    OPEN(UNIT=25,FILE='cvmlogn6.dat',STATUS='unknown')
  ENDIF
  IF(RUNTYPE.EQ.7)THEN
    OPEN(UNIT=10,FILE='adunif7.dat',STATUS='unknown')
    OPEN(UNIT=11,FILE='adbeta7.dat',STATUS='unknown')
    OPEN(UNIT=12,FILE='adexpn7.dat',STATUS='unknown')
    OPEN(UNIT=13,FILE='addblx7.dat',STATUS='unknown')
    OPEN(UNIT=14,FILE='adweib7.dat',STATUS='unknown')
    OPEN(UNIT=15,FILE='adlogn7.dat',STATUS='unknown')
    OPEN(UNIT=20,FILE='cvmunif7.dat',STATUS='unknown')
    OPEN(UNIT=21,FILE='cvmbeta7.dat',STATUS='unknown')
```

```
      OPEN(UNIT=22,FILE='cvmexpn7.dat',STATUS='unknown')
      OPEN(UNIT=23,FILE='cvmdblx7.dat',STATUS='unknown')
      OPEN(UNIT=24,FILE='cvmweib7.dat',STATUS='unknown')
      OPEN(UNIT=25,FILE='cvmlogn7.dat',STATUS='unknown')
     ENDIF

     DO 1000 N = 10,MAX,10

*********************************************************
* INITIALIZE EVERYTHING FOR THIS RUN OF SAMPLE SIZE N.  *
*********************************************************


      DO 50 DISTRIB = 1,NUMDIST
         DO 50 STATTYPE = 1,2
            DO 50 ALPHA = 1,5
               POWER(DISTRIB,STATTYPE,ALPHA) = 0.0
50    CONTINUE

      Y=0.0
      P=0.0
      ANDAR=0.0
      AND=0.0
      CRAMER1=0.0
      CRAMER2=0.0

      DO 51 I = 1,N
        S(I)=0.0
51    CONTINUE

*********************************************************
*********************************************************
*********************************************************

      DO 600 DISTRIB = 1,NUMDIST

        CALL RNSET(SEED(DISTRIB))

        DO 500 J = 1,IREPSIZE

C         GENERATE THE RANDOM DEVIATES!!!!!!
          GO TO(60,65,70,75,80,85)DISTRIB
```

```
      WRITE(*,55)
55    FORMAT('ERROR AT GENERATION COMPUTED GO TO.')
      STOP

C     UNIFORM,BETA,EXPON,DOUBL EXPON,WEIBULL, & LOGNORMAL
C     DETAILS ON THESE RANDOM # GENERATORS CAN BE HAD FROM
C     THE IMSL STAT/LIBRARY MANUAL WITH THE RESPECTIVE PAGES:
C     P963,   P993, P999, P1001,     P1025,    P1011.

60    CALL RNUN(N,S)
      GO TO 90

65    CALL RNBET(N,3.0,2.0,S)
      GO TO 90

70    CALL RNEXP(N,S)
      GO TO 90

75    CALL RNEXT(N,2.0,1.0,0.5,S)
      GO TO 90

80    CALL RNWIB(N,2.0,S)
      CALL SSCAL(N,6.0,S,1)
      GO TO 90

85    CALL RNLNL(N,0.0,1.0,S)

C     STANDARDIZE AND SORT THE STANDARDIZED DATA!!!!!!!
95    FORMAT(A15,2X,I2,/,10(F5.2,1X))
90    CALL STANDAR(S,N)
      CALL SVRGN(N,S,S)

      DO 400 L = 1,N
         Y = S(L)
C        DETERMINE THE CDF VALUES
C        FOR THE STANDARDIZED DATA.
         S(L) = ANORDF(S(L))
400   CONTINUE

C     FIND THE STAT VALUES FOR THE FOLLOWING TESTS.

      CALL ANDERSON(N,S,ANDAR,RUNTYPE)
```

```
      CALL CRAMER(N,S,CRAMER1,RUNTYPE)

C     ENTER THE CRITICAL VALUES FOR THE SAMPLE OF SIZE N FROM
C     THE CRITICAL VALUE TABLES.

      DO 450 ALPHA = 1,5

        IF (ANDAR  .GT.CRITVAL(1,ALPHA,N/10))
     +     POWER(DISTRIB,1,ALPHA) = POWER(DISTRIB,1,ALPHA) + 1.0
        IF (CRAMER1.GT.CRITVAL(2,ALPHA,N/10))
     +     POWER(DISTRIB,2,ALPHA) = POWER(DISTRIB,2,ALPHA) + 1.0

450     CONTINUE


500    CONTINUE

*****************************************************************
*                                                               *
*          DETERMINE THE REJECTION %AGE, THE POWER              *
*                                                               *
*****************************************************************

     DO 550 STATTYPE = 1,2
        DO 550 ALPHA = 1,5
          POWER(DISTRIB,STATTYPE,ALPHA)=POWER(DISTRIB,
     +                         STATTYPE,ALPHA)/REAL(IREPSIZE)
550    CONTINUE


600   CONTINUE

*****************************************************************
*                                                               *
*    PRINT THE HEADERS AT THE TOP OF ALL 12 FILES.              *
*                                                               *
*****************************************************************

      IF(N.EQ.10)THEN
        DO 605 STATTYPE = 1,2
          DO 605 DISTRIB = 1,NUMDIST
            WRITE(STATTYPE*10+DISTRIB-1,603)DISTNAM(DISTRIB),
```

```
     +              STATNAM(STATTYPE),COMBONAM
603          FORMAT(4X,'ACTUAL POPULATION: ',A20,//,
     +              2X,'STATISTIC: ',A34,//,
     +              2X,'PLOT POSITION IS ',A32,///,
     +              ,7X,'POWERS AT ALPHA LEVELS',//,
     +               'N',3X,'.20   .15   .10   .05   .01'/)
605       CONTINUE
        ENDIF

*******************************************************************
*                                                                 *
*       PRINT THE POWERS                                          *
*                                                                 *
*******************************************************************

        DO 620 STATTYPE = 1,2
          DO 620 DISTRIB = 1,NUMDIST
C           PRINT THE REJECTION %AGE, THE POWER
            WRITE(STATTYPE*10+DISTRIB-1,610)N,(POWER(DISTRIB,STATTYPE,
     +        ALPHA),ALPHA=1,5)
610         FORMAT(I2,2X,5(F5.3,3X))
620     CONTINUE

1000  CONTINUE

      DO 1100 I = 1,2
         DO 1100 J = 1,6
           CLOSE(I*10+J-1)
1100  CONTINUE

      RETURN

      END

*****************************************************************
*                                                               *
*                   THE END OF THE ROAD!!!!!!                   *
*                                                               *
*****************************************************************
```

Bibliography

1. Amstadter, Bertram L. *Reliability Mathematics*. New York: McGraw-Hill Book Company, 1971.
2. Banks, Jerry and John S. Carson. *Discrete-Event System Simulation*. Englewood Cliffs: Prentice-Hall, 1984.
3. Bradley, James V. *Distribution-Free Statistical Tests*. N.J: Prentice Hall Inc, 1968.
4. Bush, John G. *A Modified Cramer-von Mises and Anderson-Darling Test for the Weibull Distribution with Unknown Location and Scale Parameters*. MS thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1982.
5. Conover, W.J. *Practical Non-parametric Statistics*. New York: John Wiley and Sons Inc., 1971.
6. D'Agostino, R, B. "An omnibus test for normality for moderate and large sample size.," *Biometrica*, *58* (1971).
7. D'Agostino, Ralph B. and Michael A. Stephens. *Goodness-of-fit Techniques*. New York: Marcel Dekker, Inc., 1986.
8. Efron, Bradley. *The Jacknife, the Bootstrap and Other Re-sampling Plans*. Bristol, England: J. W. Arrowsmith Ltd., 1982.
9. Green, J. and Y. Hegazy. "Powerful Modified EDF Goodness-of-Fit Tests," *Journal of American Statistical Association*, *I*(I) (1976).
10. Hammersley, J. M. and D. C. Handscomb. *Monte Carlo Methods*. London: Methuen and Co., 1967.
11. Harter, H. L. "Another Look at Plotting Positions," *Communications in Statistics - Simulation*, 317–343 (1984).
12. Harter, H. L. "A Monte Carlo Study of Plotting Positions," *Communications in Statistics - Simulation*, 317–343 (1985).
13. Kahya, G. *New Modified Anderson-Darling and Cramer-von Mises Goodness-of-Fit Tests for a Normal Distribution with Specified Parameters*. MS thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1991.
14. Lapin, Lawrence L. *Statistics for Modern Business Decisions*. New York: Harcourt Brace Janovich, Inc, 1981.
15. Massey, Frank J. "The Kolmogorov-Smirnov Test for Goodness-of-Fit," *Journal of American Statistical Association*, *I*(I) (1951).
16. Mendenhall, W., et al. *Mathematical Statistics with Applications*. Boston: PWS-KENT Publishing Company, 1990.

17. Moore, A. H., "Class lecture in STAT 699, Goodness-of-Fit Tests." School of Engineering, Air Force Institute of Technology, July 1992. Wright-Patterson AFB, OH.

18. Moore, A. H. et. al. "Modified Goodness-of-fit tests for the Logistic Distribution with Unknown Location and Scale Parameters.," *Communication in Statistics - Simulation*, *I*(15) (1986).

19. Pearson, E. S. et. al. "Tests for departure from normality. Comparison of Powers.," *Biometrica*. *I*(64) (1977).

20. Porter, James E. *Modified Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises Tests for the Pareto Distribution with Unknown Location and Scale Parameters.* MS thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1985.

21. Pritsker, A. Alan B. *Introduction to Simulation and SLAM II*. New York: John Wiley and Sons, Inc., 1984.

22. Ream, Thomas J. *A New Goodness-of-Fit Test for Normality with Mean and Variance Unknown*. MS thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1981.

23. Stephens, M. A. "Use of the Kolmogorov-Smirnov, Cramer-von mises and related statistics without extensive tables.," *Journal of the Royal Statistical Society*, *I*(B32) (1970).

24. Viviano, Philip J. *Modified Kolmogorov-Smirnov, Anderson-Darling, and Cramer-von Mises Tests for the Gamma Distribution with Unknown Location and Scale Parameters*. MS thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1982.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE March 1993	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE
MODIFIED ANDERSON-DARLING AND CRAMER-VON MISES GOODNESS-OF-FIT TESTS FOR THE NORMAL DISTRIBUTION

5. FUNDING NUMBERS

6. AUTHOR(S)
David A. Gwinn, Sr., Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GOR/ENS/93M-07

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

New techniques for calculating goodness-of-fit statistics for normal distributions with parameters estimated from the sample are investigated. Samples are generated for a Normal(0,1) distribution. Critical values are calculated for five modifications to the Anderson-Darling statistic and five modifications to the Cramer-Von Mises statistic. An extensive power study is done to test the power of the new statistics versus the power of the unmodified statistics.
Powers of six of the new statistics show minimal to no improvement, two of the new statistics show a marked decrease in power, and two of the new statistics show an overall increase in power over the unmodified statistics. One of these two improved statistics was the obvious better of the two, and it was a modification to the Anderson-Darling statistic.
Complete tables of critical values for sample sizes n=4 through n=50 are included for all Anderson-Darling and Cramer-Von Mises statistics, both modified and unmodified.

14. SUBJECT TERMS
Normal Distribution; Goodness-of-Fit; Monte Carlo Simulation; Anderson-Darling; Cramer-von Mises

15. NUMBER OF PAGES
92

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL